DESIGN, OPTIMIZATION AND SYNTHESES OF SMALL MOLECULES TO DISRUPT PROTEIN-PROTEIN INTERACTIONS

A Dissertation

by

JARU TAECHALERTPAISARN

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Kevin Burgess
Committee Members,	Frank M. Raushel
	Coran M. H. Watanabe
	Thomas R. Ioerger
Head of Department,	Simon W. North

May 2018

Major Subject: Chemistry

Copyright 2018 Jaru Taechalertpaisarn

ABSTRACT

Protein-protein interactions (PPIs) are one of the basic mechanisms in cellular biology, but also involve in diseases if they are dysregulation. Disrupting aberrant PPI activities is useful in medicinal chemistry. One approach to inhibit PPIs is to design small molecule peptidomimetics bearing side-chain orientations similar to protein ligands, in which those mimics might displace or interfere the native PPIs. Previous research in our group developed Exploring Key Orientations (EKO) program that matches $C\alpha$-$C\beta$ coordinates of virtual small molecules to the side-chain vectors of proteins at PPI interfaces. Similar $C\alpha$-$C\beta$ orientations between mimics and protein ligands indicate that small molecules might be suitable to displace protein ligands, i.e. those compounds might interfere PPIs.

We used EKO to deduce small molecules that might disrupt medicinally-relevant PPIs. Herein, EKO implicated our designed mimics, hydantoin-oxazoline, triazole-oxazole and triazole-oxazoline derivatives, might disrupt Nef•MHC-I•AP1 and NEDD8•NAE interactions, in which they are relevant to HIV-1 and cancer diseases respectively. After learning from these projects, we designed hydantoin-piperazine analogues to disrupt PCSK9•LDLR interaction that causes hypercholesterolemia disease. Although the first-generation hydantoin-piperazine derivatives did not show good PCSK9•LDLR inhibition, we modified chemotype structures by cooperating with a docking program, Glide, to improve inhibitory potencies. As a result, we successfully obtained lead compounds that significantly disrupt PCSK9•LDLR interaction with the measurable binding affinities. Besides these protein targets, we synthesized another minimalist mimic, oxazoline piperidine-2,4-dione, that has conformational biases toward helical and sheet-turn-sheet

motifs. This structure potentially has favorable cellular- and oral-permeability calculated by QikProp.

We are also interested in how to design molecules suitable for PPI inhibition. A concept of secondary structure mimicry is widely applied to design molecules that resemble a secondary structure at an PPI interface, hence possibly disrupt protein-protein interaction. However, there is no direct study to prove a correlation between secondary structure mimicry and interface mimicry. To respond this issue, we used EKO to match several new chemotypes on the ideal secondary structures and PPIs database, and then compared the frequencies of secondary structures that chemotypes matched at PPI interfaces to the ideal secondary structure biases of each chemotype. We found that, in general, good secondary structure mimics tend to match frequently at PPI interfaces; however, they mostly match on non-ideal secondary structure motifs.

ACKNOWLEDGEMENTS

I would like to express my gratitude to my advisor, Dr. Kevin Burgess, for his guidance and support throughout the course of this research. I also thank my committee members, Dr. Frank Raushel, Dr. Coran Watanabe, and Dr. Thomas Ioerger for their guidance.

I would like to acknowledge Dr. Lisa Perez for her teaching and troubleshooting all computational works, Dr. Yong Xiong for testing my Nef•MHC-I•AP1 inhibitors, Dr. Jennifer Carew for testing my NEDD8•NAE inhibitors, Dr. Jay Horton for his kindness to provide PCSK9 protein for my research, Dr. Xiaowen Liang for her effort to do SPR study for my PCSK9 project, Bosheng Zhou for his photoaffinity labeling work, and Zhengyang Jiang for teaching and consulting all cell-based works.

I would like to thank to all former Dr.Burgess's members, Dr. Eunhwa Ko, Dr. Sakunchai Khumsabdee, Dr. Anyanee Kamkaew for their advice, and special thanks to Dr. Dongyue Xin for his advice and support. Also, I would like to thank all the current members for their companionship.

I would like to thank Jill Powers for her assistance since I joined Dr. Burgess's group until her retirement.

Finally, I would like to thank my family for their love, patience and support.

CONTRIBUTORS AND FUNDING SOURCES

This work was supported by a dissertation committee consisting of Professors Kevin Burgess [advisor], Frank Raushel, and Coran Watanabe of the Department of Chemistry and Professor Thomas Ioerger of the Department of Computer Science.

The inhibition experiments of Nef•MHC-I•AP1 and NEDD8•NAE in Chapter II were conducted by Professor Yong Xiong of the Department of Molecular Biophysics and Biochemistry at Yale University, and Professor Jennifer Carew of the University of Arizona Cancer Center respectively. SPR experiments and analyses in Chapter IV were conducted by Dr. Xiaowen Liang from the Center for Infectious and Inflammatory Diseases, Institute of Biosciences and Technology. Photoaffinity labeling in Chapter IV was done by Bosheng Zhou from Dr. Burgess's group. Flow cytometer was operated by Zhengyang Jiang from Dr. Burgess's group. Data analyses in Chapter V were done in part with Maritess Arancillo, Dr. Cheng-Ming Lin, Rui-Liang Lyu, and Zhengyang Jiang from Dr. Burgess's group. Compounds syntheses in Chapter VI were performed in part with Dr. Xun Li and Dr. Dongyue Xin from Dr. Burgess's group. All other works, including computational studies, all syntheses and characterizations in Chapter II-V, binding and biological studies in Chapter III and IV, data analyses in Chapter V and VI, were completed independently by the student.

Graduate study was supported by The Queen Sirikit Scholarship under The Royal Patronage of Her Majesty Queen Sirikit of Thailand, and by a fellowship from Department of Chemistry, Texas A&M University.

TABLE OF CONTENTS

Page

Page

Page

LIST OF FIGURES

Page

Page

Page

Page

Page

LIST OF TABLES

Page

LIST OF SCHEMES

Page

CHAPTER I

INTRODUCTION

1.1 Disrupting Protein-protein Interactions By Small Molecules

Protein-protein interactions (PPIs) are pivotal to cellular biology. They are implicated in signal transduction, cell-to-cell communication, transcription, replication and membrane transportation. Dysregulation of PPIs can cause neurodegenerative diseases, cancer, and other maladies.[1] Inhibitors that disrupt aberrant PPIs associated with diseases states are potentially useful in medicinal chemistry. Theoretically, this type of inhibitor should have an advantage over enzyme inhibitors, for instance, insofar as they selectively bind on protein targets due to the diversity of protein interfaces unlike targeting similar catalytic active sites of enzymes; this reduces probability of off-target effects. However, protein interfaces tend not to have well-defined, concave binding sites and most are lipophilic, thus hard to 'drug' using small molecules.[2,3] Even though protein interfaces are featureless, they also contain 'hot-spot' residues, which are largely responsible for driving binding.[1,4] Hot spots tend to tightly pack in clusters forming a network of interactions, in which can be called 'hot-regions'.[5] Interactions around 'hot regions' are dominated by side-chain-to-side-chain contacts, which are accounted for ~80% of all interfacial interactions.[6] Thus small molecules that perturb 'hot-spot' or 'hot-regions' binding sites are favorable candidates to disrupt PPIs.

Several methods are currently used to discover PPI inhibitors. High-throughput screening (HTS) is a widespread approach; this suffers low hit rates, weakly potent hits and difficulties in removal of false positives because compound collections do not have

appropriate chemotype.[1,7] Fragment-based screening features elaboration of low-affinity small molecules that are not likely to disrupt PPIs alone;[8] the challenge with this method tends to be linking fragments together effectively. Computer-aided design approaches tend to be quite limited[9]. Of particular interest is rational design of secondary structure mimics[10-12] following the hypothesis that if a secondary structure exists at an interface, then a mimic that presents side-chains in the same way might displace the protein ligand in the PPI.

A widely applied design principle for designing molecules to inhibit PPIs is the concept of *secondary structure mimicry*.[11,12] In applying this method, researchers identify a secondary structure at an interface, possibly one known to encompass a hot-spot,[13] then focus on designing molecules that resemble that secondary structure. Some chemotypes featured in this method are peptidic. For instance, stapled peptides[14-22] which are essentially cyclic peptides, have been used by many groups to mimic helices at interfaces. However, there are also considerable efforts to use less peptidic molecules that do not have polyamide backbones, but do present amino acid side-chains.[23-26] These have come to be known as *minimalist secondary structure mimics*.

1.1.1 Minimalist Peptidomimetics

Minimalist mimics do not have peptidic backbones; instead, they have rigid organic cores that can project at least three amino acid side-chains in orientations that resemble similar features on select regions of peptides or proteins. Examples of minimalist mimics are shown in Figure 1.1 (putative helix mimics: **A**,[27-29] **B**,[30] **C**,[31-33] **D**,[34,35] **E**,[36] **F**,[37] **G**,[38] and **H**[39]; sheet mimic, **I**[40]). Much research in this area has been driven by the assumption

that if a secondary structure is found at a protein-protein interface then a corresponding minimalist mimic is a candidate to displace the protein containing that motif from the interface.[10,11,26] There are, in fact, several instances in which minimalist mimics have been proven to disrupt PPIs.[40-51] For example, terphenyl-based helical mimetics (**B**) was proved to disrupt p53/HDM2 involving cancerous tumor.

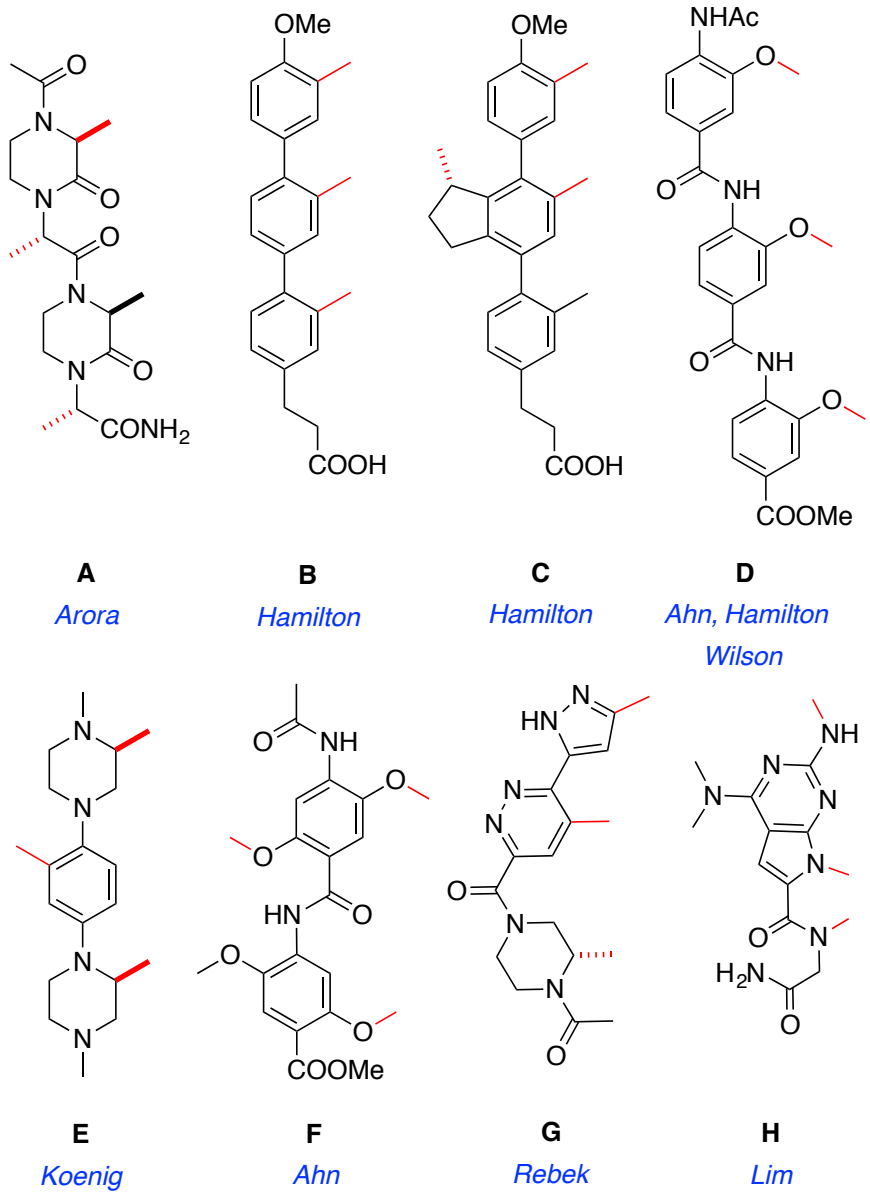
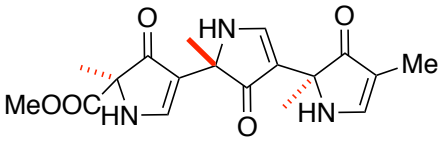


Figure 1.1 Examples of minimalist mimics **A**-**I**.



I

Smith

Figure 1.1 Continued.

According to the description outlined above, several compounds can be claimed as the minimalist peptidomimetics. However, what are the criteria of being good minimalist mimics? In our opinion, there are four characteristics that guide researchers to design good minimalist mimics.[52]

(i) Syntheses of chemotypes must be facile for most of amino acid side-chains (e.g. Arg, Trp, His etc.)

(ii) Minimalist mimics may not have global minimum energy conformations corresponding to the proteins or peptides they resemble, but those corresponding conformations should be thermodynamically or kinetically accessible at given conditions (e.g. at room temperature or physiological temperature).

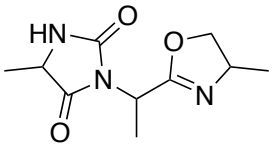
(iii) Chemotypes should be rigid to compensate the loss of entropy on docking.

(iv) Accessible chemotypes project *C*α-*C*β coordinates that overlay well on the *C*α-*C*β of secondary structures or proteins.

1.1.2 Universal Peptidomimetics

Secondary structures at PPI interfaces are most often distorted, and many interfaces do not involve them at all. Thus, any mimic that exclusively resembles an *ideal* secondary

structure motif is unlikely to be widely applicable to interfaces containing real secondary structures of that kind. The hidden potential of so-called minimalist mimics is that many chemotypes in this class also populate conformations other than the targeted secondary structure (Figure 1.2).[52-54] We coined the term *universal peptidomimics* to describe small molecules simultaneously populating conformations that resemble more than one secondary structure.[53] This led us to appreciate that, in some situations, minimalist mimics can resemble regions of a PPI interface without resembling *any* particular secondary structure. Consequently, it is important to directly relate accessible conformations of minimalist mimics to PPI interfaces without reference to secondary structures.



1

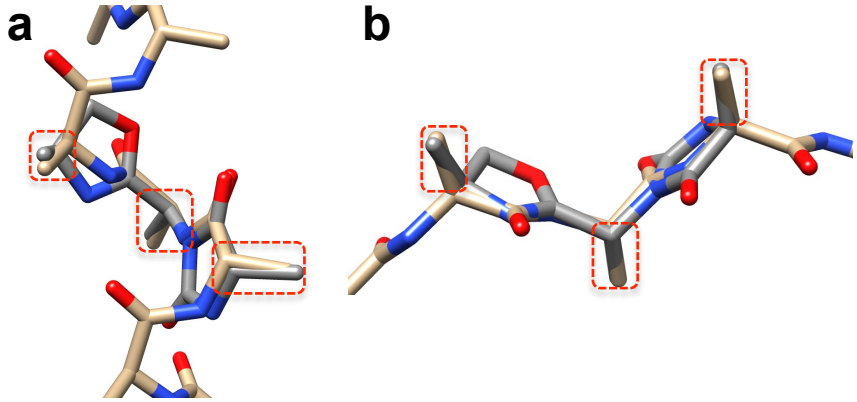


Figure 1.2 An example of a universal peptidomimetic. Compound **1** can adopt their conformations to match on (**a**) 3_{10}-helix, (**b**) β-strand, (**c**) α-helix, (**d**) strand-turn-strand, and (**e**) non-ideal secondary structure at PPI interface.

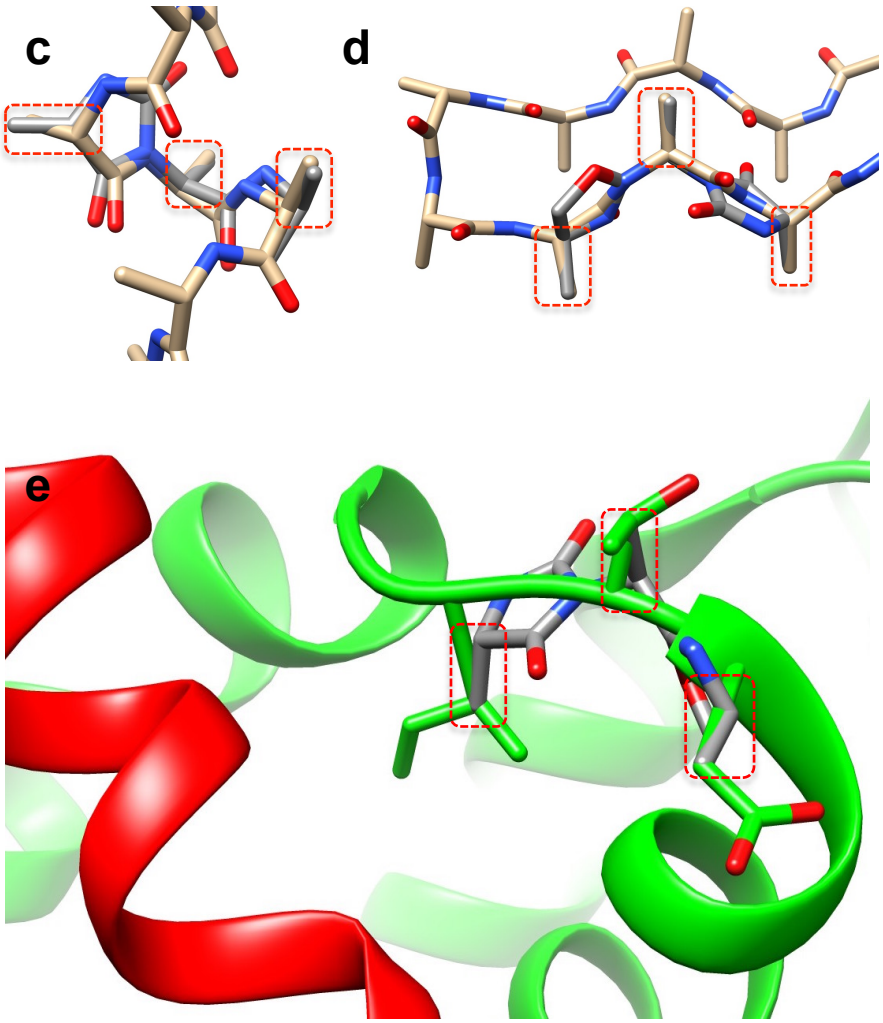


Figure 1.2 Continued.

1.2 Exploring Key Orientation (EKO) as a Method to Identify Small Molecules that Disrupt Protein-protein Interaction

1.2.1 Introduction

Minimalist peptidomimetics have to project side-chains vectors similar to the side-chains orientations of protein ligands to mimic side-chains interactions between protein

receptors and ligands. Nevertheless, the ability to obtain good $C\alpha$-$C\beta$ overlays is beyond human perception. This is because rigid chemotypes can adopt several conformations in solution. Identifying conformations by crystallography may be inappropriate since this method reveals only one conformation, and that crystal structure may not be a good representative due to the distortion by crystal packing forces. Dr. Burgess' lab had developed algorithms to systematically overlay side-chains of numerous conformations of a chemotype to the side-chains of PPIs or secondary structures. We named this program 'Exploring Key Orientation (EKO)'. Figure 1.3 illustrates steps in this strategy. This program enumerates simulated preferred conformations using quenched molecular dynamics (QMD, described below). $C\alpha$-$C\beta$ coordinates of conformers are clustered, overlaid on the $C\alpha$-$C\beta$ coordinates of PPIs or secondary structures, and determined how goodness of fit in root mean square deviation (RMSD). Low RMSDs imply the side-chains orientations of preferred conformations replicate side-chains orientations of PPIs or secondary structures. Side-chain vectors beyond $C\alpha$-$C\beta$ (e.g. $C\beta$-$C\gamma$) are not considered because those are relatively mobile and do not define overall direction. This program can be applied to search good overlays of chemotypes on secondary structures (EKOS), side-chains at PPI interfaces (EKO), and single crystals of chemotypes (EKOX).

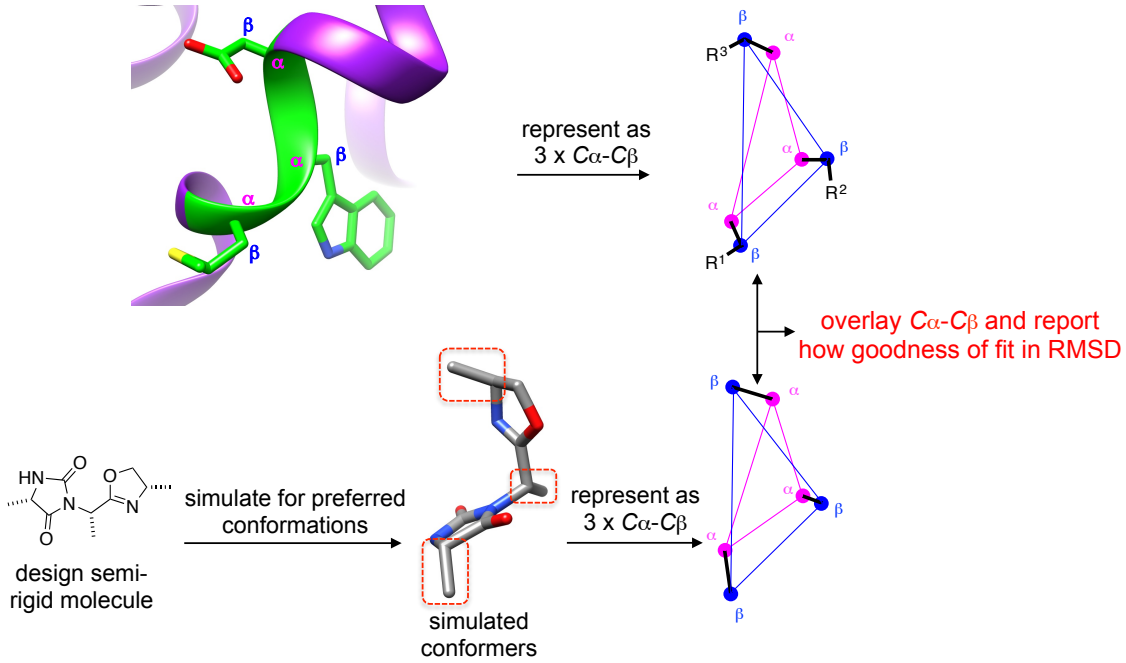


Figure 1.3 Diagram of EKO process.

1.2.2 EKOS

EKOS (*Exploring Key Orientations on Secondary structures*) facilitates systematic, quantitative overlay of minimalist mimic preferred conformations on secondary structures (Figure 1.4).[54,55] This procedure can be applied to ideal secondary structures, interfacial secondary structures (that may be distorted from the ideal), or any protein fragment for which there is crystallographic data or reliable PDB coordinates.

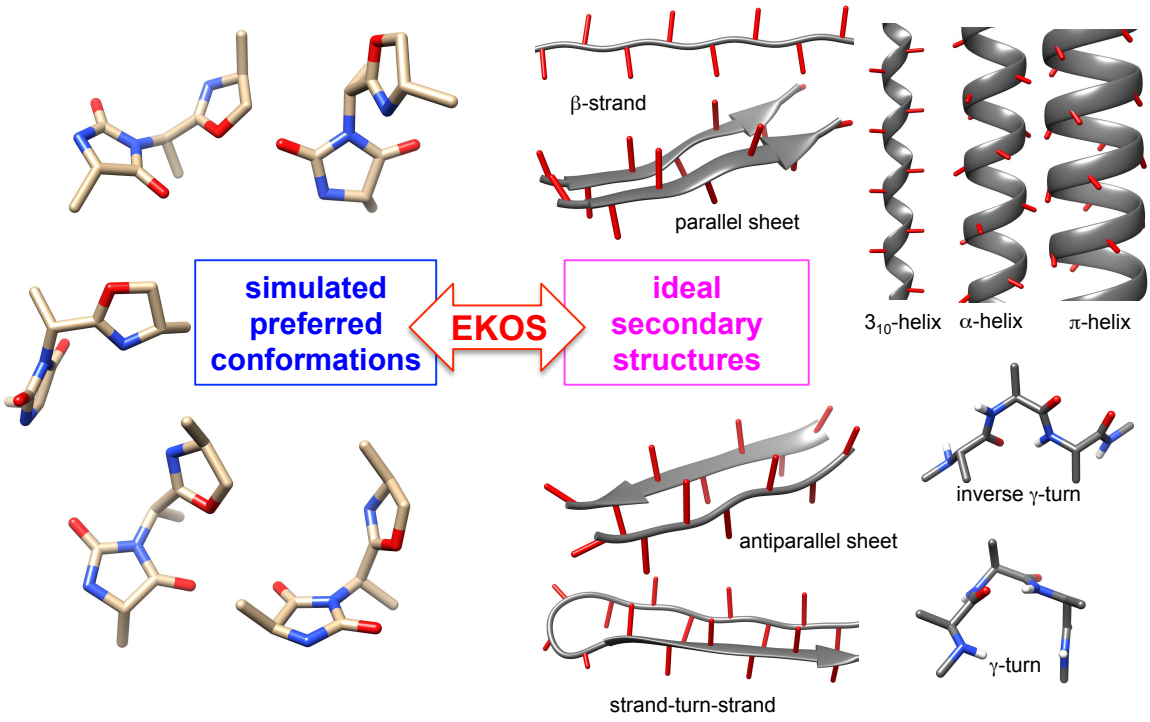


Figure 1.4 Diagram of EKOS

In EKOS, preferred conformations of the featured minimalist mimic with three methyl side chains (representing Ala, Ala, Ala) are determined using a molecular dynamics routine. Quenched molecular dynamics (QMD)[56,57] is the protocol used in our work. In QMD, a large number (e.g. 1500) of high temperature conformations for each stereoisomer of a chemotype are rapidly (no slow cooling) quenched to give conformations lacking bond vibrations (*ie* at 0 K). In each case, conformers within 3 kcal•mol^{-1} of the lowest energy one identified are described as "preferred". These preferred conformations are systematically overlaid with each other using an algorithm that optimizes fit of side-chain *C*α-*C*β vectors to generate families of conformers. Rigid chemotypes may converge to only one conformational cluster, whereas hundreds of families might be generated from more flexible systems. A key feature of the families is that every member projects side-chains in highly similar directions, since they are

clustered on the basis of $C\alpha$-$C\beta$ vectors. Families generated in QMD at this stage are similar to dynamic Boltzmann distributions of conformational states in solution.

QMD implemented as described above is deliberately designed to reveal the intrinsic conformational biases of a chemotype in the absence of any extrinsic factors. It is for this reason that the calculations are performed using a continuous dielectric medium of 80 (roughly approximating to water), and not using explicit solvent molecules. Similarly, three methyl substituents are used so that the conformations of the chemotype are not biased by interactions between side-chains.

1.2.3 EKO

Another data-mining approach (*Exploring Key Orientations, EKO*) similarly compares simulated preferred conformations of minimalist mimics but with sets of three *interface side-chains in PPIs* (Figure 1.5).[58] Intrinsic conformational biases of chemotypes are critical because if a small molecule tends to project side chains in the same orientations as a segment of the protein ligand in a PPI, then that protein receptor can at least accommodate, and perhaps even enforce, this conformational bias.

EKO can be run in two ways; we call these *matching* and *mining*, and they are biology- and chemistry-centered, respectively (Figure 1.5). In EKO *matching*, a single protein-protein interface is compared with preferred conformations of a series of chemotypes. This is *biology-centered* because the PPI target is decided first, supposing that a chemotype can then be found to disrupt it. This approach is in perfect harmony with

common strategies in the pharmaceutical industry in which target selection is the first step.

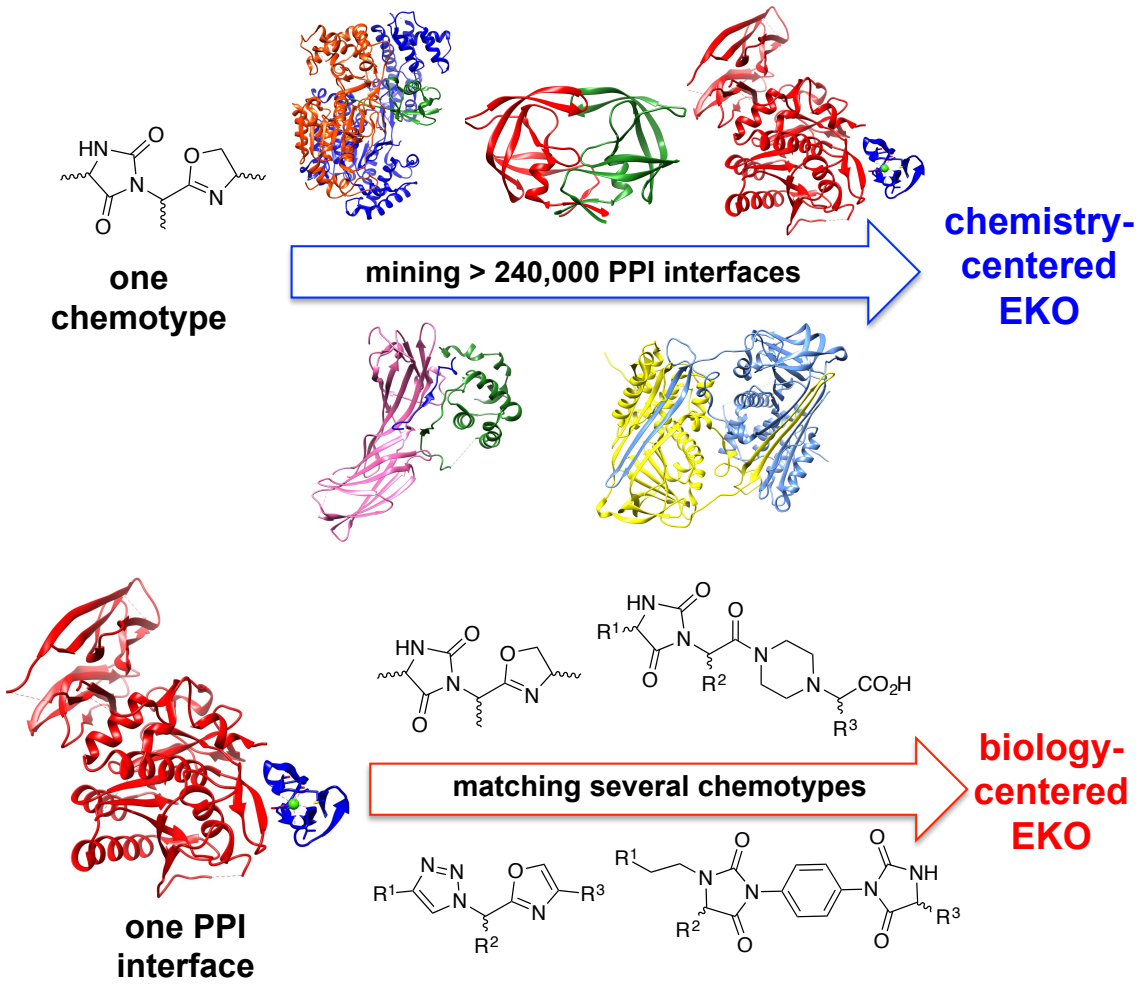


Figure 1.5 Two modes of EKO: chemistry-centered EKO (mining, top), and biology-centered EKO (matching, bottom).

EKO *mining* compares preferred conformations of one stereoisomer of a chemotype with nearly all the PPI interfaces in the PDB (currently >240,000, greater than the number of structures in the PDB). Using queued jobs, eight stereoisomers of a chemotype can be compared in this way in about 8 h using a powerful desktop computer with ample disc space. A chemotype that mimics peptide/protein conformations well, in terms of side-chain orientations, will inevitably overlay on medicinally relevant PPIs. Researchers in

the pharmaceutical industry tend to have less interest in EKO mining than the matching routine. This comes from a valid observation that target selection is such an important issue, but perhaps an under-appreciation of how difficult it is to design small molecules to impede specific PPIs. EKO mining has the advantage that using it to identify potentially tractable PPIs for a chemotype is a simpler problem than *de novo* chemotype design for a specific PPI. At a minimum, EKO mining is an idea-generation tool to narrow the target selection down to PPIs that are compatible with the minimalist mimics that can be made.

1.2.4 EKOX

Crystal structures of minimalist mimics are informative, but they only give one snap-shot of the dynamic conformational ensemble. There is no widely used method to evaluate if a conformation observed in a single crystal X-ray study is a preferred conformation in a dynamic equilibrium, or just an artifact of crystal packing forces. We believed it was desirable to be able to quantify the correspondence between solid-state structures and conformations in solution. However, a common problem in conformational analyses is to evaluate the relevance of a solid-state structure to a conformational ensemble in solution. For minimalist mimics, a simple modification of EKOS, EKOX (*Exploring Key Orientations on X-ray structures*) can be used to determine if a solid-state conformation from an X-ray structure is present in the simulated conformational ensemble. EKOX facilitates estimates of the energy difference between the simulated solution-state conformer closest to the solid-state structure and the lowest energy one overall.

Figure 1.6 shows data to illustrate how EKOX can work. Coordinates for a solid-state structure of L-**2fff** were obtained from single crystal X-ray diffraction. Preferred solution conformers of L-**2aaa** were generated, and their *C*α - *C*β coordinates systematically compared with the X-ray structure (L-**2fff**). Figure 1.6b shows the *C*α - *C*β coordinates of the preferred conformer for L-**2aaa** in solution matched the X-ray structure of L-**2fff** with an RMSD of 0.09 Å, and that conformer was 0.16 kcal•mol^{-1} above the lowest energy one observed. Thus, the solid-state structure is a representative of at least one populated conformer in solution, and is not an artifact of crystal packing forces.

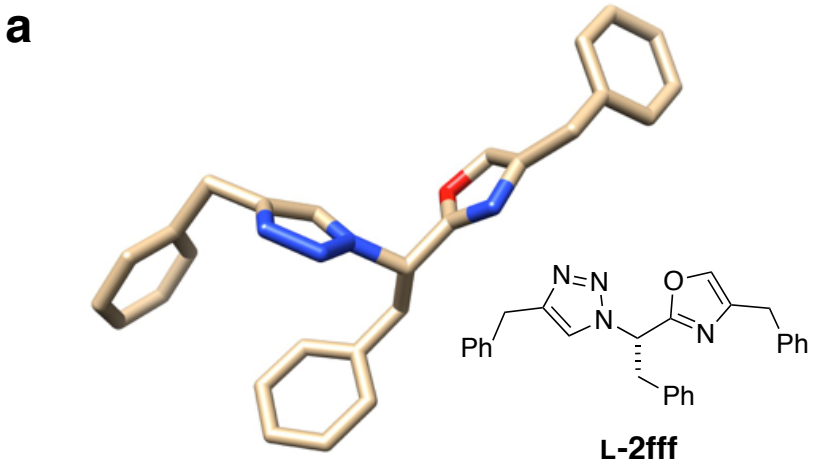


Figure 1.6 Illustration of EKOX. (**a**) crystal structure of L-**2fff**. (**b**) Comparison of best fitted preferred conformers of minimalist mimic L-**2aaa** (grey) with the X-ray crystal structure of L-**2fff** (gold) to elucidate the match and energy difference.

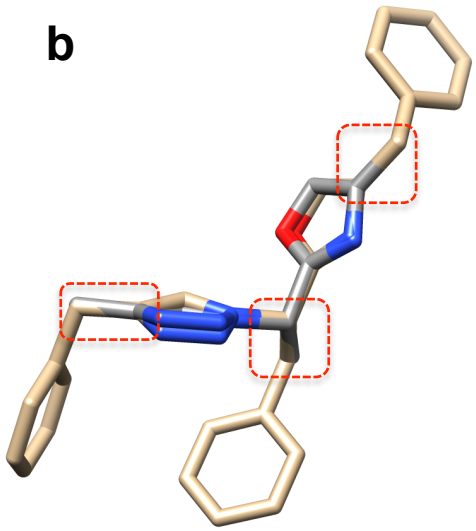


Figure 1.6 Continued.

1.3 Conclusion

Finding suitable inhibitors to disrupt aberrant PPIs involving diseases is a bottleneck in medicinal chemistry. Unlike conventional well-defined targets like kinases active sites or G-protein coupled receptors, protein-protein interfaces tend to be flat, featureless, and difficult to target. Established methods like HTS are not suitable to identify candidates within reasonable time, and tend to have high cost overheads. New methods that help accelerate the discovery of PPI inhibitors are potentially valuable.

Rational small molecules design to disrupt PPIs usually concentrates on mimicking secondary structures found at protein interfaces. Researchers attempt to synthesize less peptidic molecules bearing side-chain residues; this type of compounds is called 'minimalist peptidomimetic'. Nevertheless, comparing $C\alpha$-$C\beta$ overlays between minimalist mimics and secondary structures is difficult because chemotypes can adopt several conformational ensembles in solution. Previous works in Dr. Burgess' group

developed EKO to systematically overlay $C\alpha$-$C\beta$ coordinates of simulated preferred conformations on the sets of $C\alpha$-$C\beta$ coordinates of secondary structures or PPI interfaces. This algorithm matches preferred conformations of chemotypes on ideal secondary structures (EKOS), non-ideal secondary structures at PPIs interfaces (EKO), or a single crystal (EKOX). All these methods are useful to deduce suitable chemotypes to the appropriate protein targets. These approaches have been proved to find new inhibitors that disrupt HIV-1 protease dimerization and perturb antithrombin oligomerization.[58,59]

My PhD research began by using EKO to search potential inhibitors to disrupt Nef•MHC-I•AP1 (relevant to HIV) and NEDD8•NAE (for cancer) complexes. After learning from these projects, we moved to another new interesting target PCSK9•LDLR (related to hypercholesterolemia). Incorporating with a docking program to refine structures of chemotypes, we found the potential leads that inhibited PCSK9•LDLR interaction. Our interest broadened to study relationship between EKOS and EKO in term of how relevant chemotype bias toward ideal secondary structures to the frequency of secondary structures found at PPIs. Finally, we had prepared another minimalist mimic that might be useful to disrupt PPIs.

CHAPTER II

INITIAL ATTEMPTS TO USE EKO TO DESIGN INHIBITORS FOR TWO NON-IDEAL PROTEIN-PROTEIN INTERACTION TARGETS

2.1 Introduction

EKO implicates protein targets by determining how closely the side-chains on a scaffold align with side-chains at protein interface. If the fit is good, this indicates the molecule may displace the protein ligand and perturb that particular protein-protein interaction. This is a hypothesis that needs to be verified experimentally; our group's long-term objective is to understand how widely applicable this approach is. My initial attempts to do this prioritized the PPIs Nef•MHC-I•AP1 (in HIV) and NEDD8•NAE (for cancer). We were unable to find effective small molecules for these in the time available. That was not a failure of the approach but a failure in how we implemented the approach. This chapter is to elaborate on that, and to explain what we learned from this exercise.

The chemotype designs used for the two targets mentioned above were based on oxazoline and oxazole moieties coupled with hydantoin or triazole analogues (mimics **1**-**3**). We synthesized compounds implicated by EKO; but those chemotypes were so difficult to prepare that it was not possible to obtain sufficient numbers of samples in the time available. That meant, for instance, the optimization to improve the binding

interaction was not practical. Consequently, we were unable to find a promising compound lead that had sufficient measurable inhibitory activity.

Another reason we could not find inhibitors disrupting these PPIs because they both had large interface surfaces and tight binding complexes that should be problematic for small molecules to disrupt. It could be that EKO would never lead to effective chemotypes for these types of target, but we were simply not able to reach a conclusion because of the chemotype-synthesis issue. Nevertheless, the experience gained from these efforts let to positive results in the following chapters.

2.1.1 Oxazolines/oxazoles: Inspired Natural Building Blocks

Natural products can inspire design of novel chemotypes having biological activities. In marine organisms, oxazoline and oxazole moieties can be in post-translationally modified cyclic peptides.[60,61] These structures can resist enzymatic degradation, show better bioavailability than linear peptides, and enhanced conformational rigidity. Many oxazoline/oxazole-containing natural products are antibacterial, antiviral or antitumor agents.[61] Consequently, medicinal chemists frequently use derivatives of these heterocycles.

Oxazoline and oxazole scaffolds can be implemented into peptidomimetic designs as amide isosteres to prevent proteolysis[62], or as privileged pharmacophores.[63-65] Moreover, oxazoline and oxazoles can be synthesized from amino acids enabling opportunities to incorporate most of the 19 genetically encoded side-chains.

2.1.2 Nef (Negative Regulatory Factor) Protein: A Key Player In HIV

Human Immunodeficiency Virus (HIV) is a lentivirus that impairs the immune system in the host cells by reducing levels of $CD4^+$ helper T cells.[66-68] Several enzymes (*eg* reverse transcriptase, integrase or protease) and accessory proteins (*eg* Nef) have been studied intensively to cure or repress HIV infection.[69-74] Consequently there are now several drugs that target proteins associated with HIV-1. These pharmaceuticals are normally used in combination, as antiretroviral "drug cocktails". These drug combinations are so effective that they overcome the high rate of mutation in the HIV-virus. However, to eliminate the virus altogether requires a strategy to cripple it in the latent state. As a result, developments of new anti-HIV compounds are still crucial to care of AIDS patients.

Although anti-retroviral drug cocktails have been used widely to increase life expectancy for infected individuals, the rise of multidrug-resistant strains of HIV-1, and the need to overcome latency in search of a cure, underline the need for new antiretroviral drugs that have complementary mechanisms to the traditional ones. Nef inhibitors are good candidates because they interfere with the HIV-1 life cycle, and enhance recognition by immune system to kill virally infected cells (described below).[75] Within past 10 years, some inhibitors have been reported to disrupt Nef's function, either by abolishing kinase downstream signaling[76-78] or interfering Nef dimerization.[79-81] Nevertheless, an inhibitor that directly perturb tertiary complex of Nef•MHC-I•AP1 has not been reported to date. Such a small molecule that has a different mechanism of action should be considered as a complement in the antiretroviral drug cocktail.

Nef is a 23- to 35-kDa myristoylated peripheral membrane protein encoded in the HIV genome. This accessory protein enables the HIV virus to evade immune surveillance in the host by antagonizing the antiviral activity of cytotoxic T lymphocytes.[82] These T cells use MHC-I (Major Histocompatibility Complex class I) to recognize viral peptides displayed on infected cells. To avoid detection, Nef reduces the amount of MHC-I expression on cytotoxic T cells by complexing with the clathrin adapter protein complex 1 (AP1) forming a tertiary complex (Nef•MHC-I•AP1) that is internalized by clathrin-coated vesicles, wherein all components are degraded in the lysosome.[83]

The molecular basis of Nef•MHC-I•AP1 had been revealed by Guatelli and Xiong who successfully co-crystallized this complex by fusing MHC-I to the *N*-terminus of Nef and forming a single crystal with the μ1 domain of AP1 (Figure 2.1a).[84] This crystal structure revealed a binding interface of three components of around 2,266 $Å^2$. The tyrosine-containing sequence of MHC-I (YSQA) binds the canonical Yxxϕ recognition site of AP1 (ϕ, bulky hydrophobic residue; x, any amino acids) by fitting Tyr-residues of MHC-I in a binding pocket of AP1.[85] The small methyl side chain of alanine does not totally fill an AP1 pocket hence this complex does not form unless Nef is present to contribute side-chains that complete the tight association of MHC-I and AP1 (Figure 2.1c).[84] Nef uses proline-rich PxxP repeats to lock MHC-I onto the side wall of AP1. The Nef•MHC-I•AP1 interaction has two strong electrostatic interactions. First, Nef acidic cluster 62-EEEE-65 has a long-range electrostatic interaction with lysine and arginine residues of AP1 (Figure 2.1d). The second charged interface involves a three-way network from Asp327 of MHC-I, Asp123 of Nef and a basic patch of AP1 (Arg225, Arg393, Lys396, Arg211 and Arg246). Besides those binding forces, Nef *C*-terminus has an indispensable role to form a groove that binds MHC-I and AP1 (Figure 2.1b).

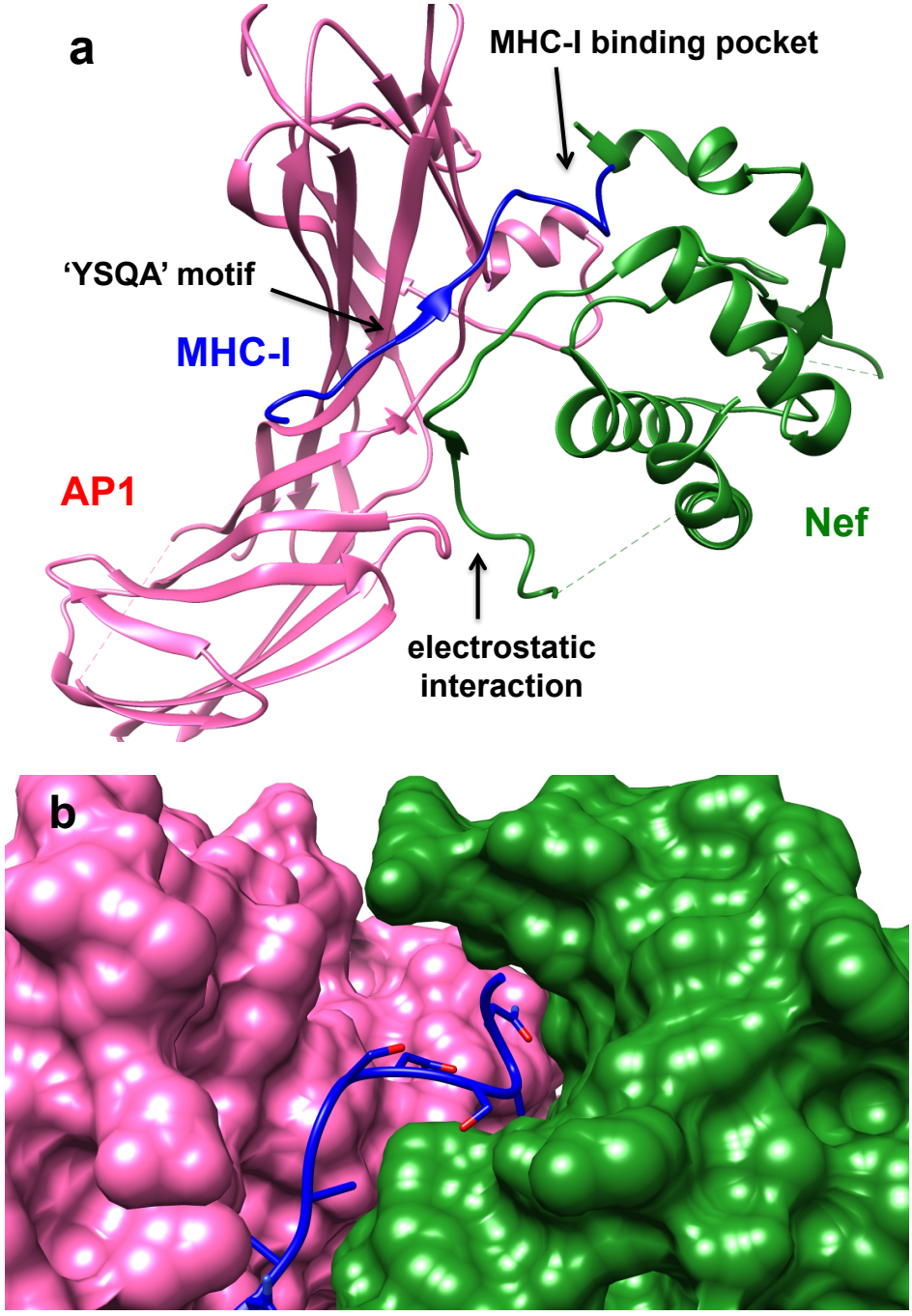


Figure 2.1 (**a**) Crystal structure of Nef•MHC-I•AP1 (4EMZ). (**b**) MHC-I binding pocket. (**c**) YSQA motif of MHC-I. (**d**) Long-range electrostatic interaction

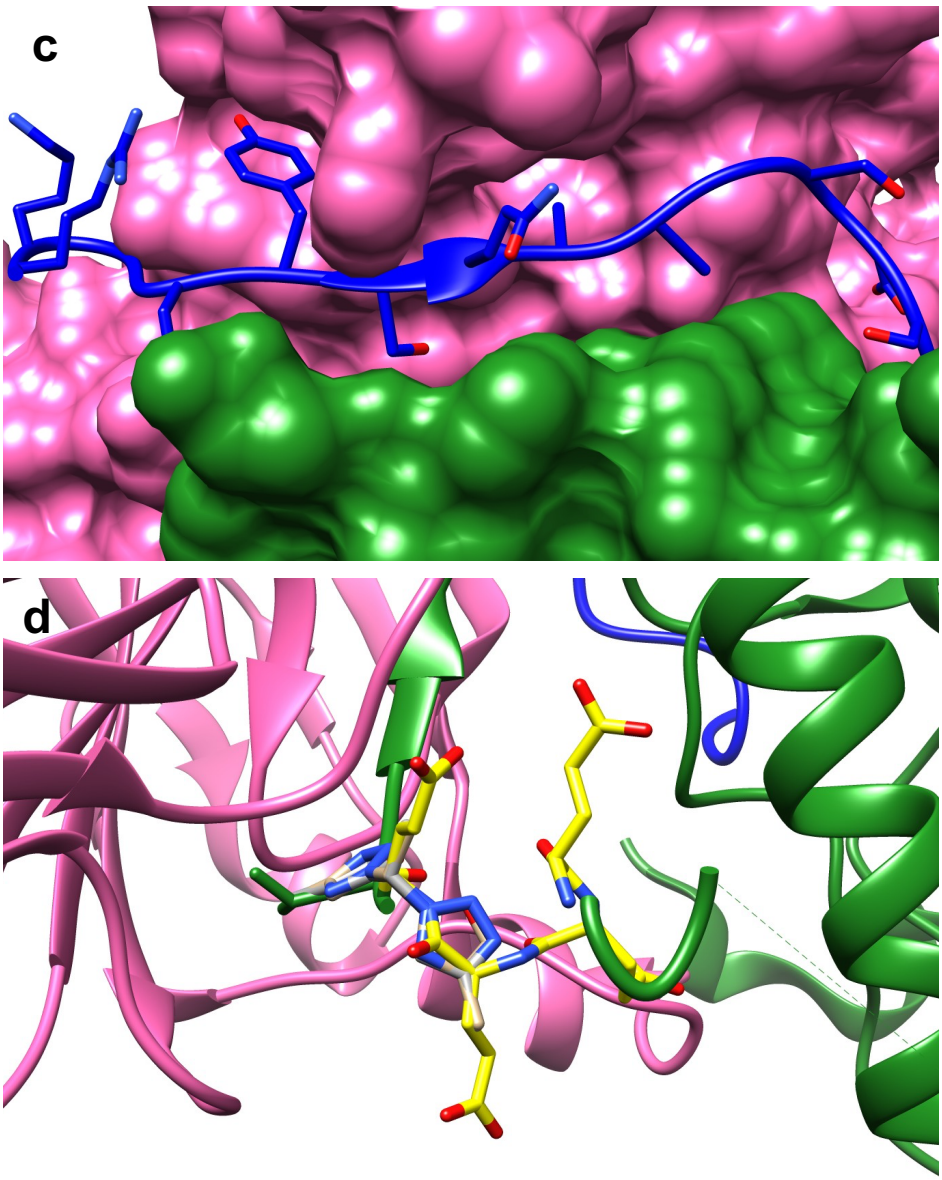


Figure 2.1 Continued.

2.1.3 NEDD8 (neural precursor cell expressed developmentally downregulated protein 8):NEDD8 ACTIVATING ENZYME (NAE)

Protein homeostasis by ubiquitination is important to cell function, and dysregulation of this process may lead to aberrant cellular activities involving cancer. The ubiquitin-

proteasome system has been validated as a therapeutic target. Thus the proteasome inhibitor, *e.g.* bortezomib (VELCADE, Millennium Pharmaceuticals Inc.), is used to treat multiple myeloma and relapsed mantle cell lymphoma.[86] However, this broad spectrum of proteasome inhibitor has some undesired off-target side effects. Therefore, inhibitors targeting at certain step, preferably E3 ligase, would selectively stabilize protein homeostasis, reduce the toxicity and improve the drug's therapeutic index.

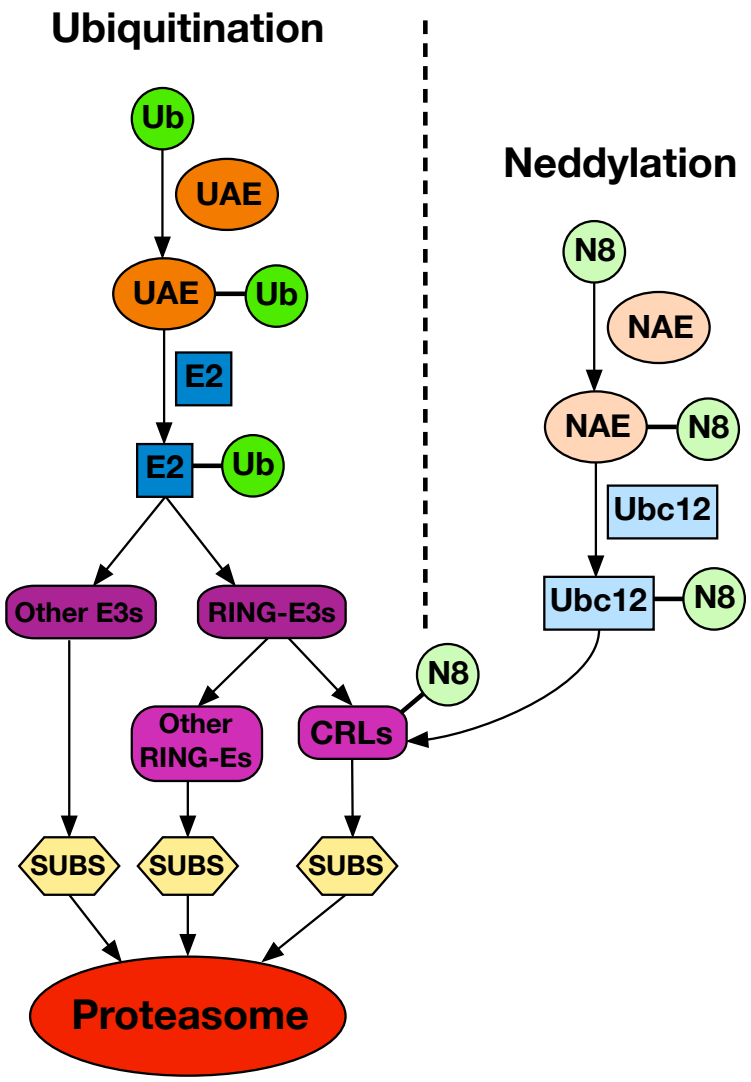


Figure 2.2 Ubiquitination and Neddylation pathway.[87]

Ubiquitinylation is regulated by a parallel mechanism involving NEDD8 and "neddylation". NEDD8 is an 81-amino acid ubiquitin-like protein with a 9 kDa molecular mass. It shares 60% identity and 80% homology with ubiquitin.[88,89] This protein specifically regulates one category of ubiquitin's E3 ligases, cullin-RING ligases (CRLs). Homeostasis of certain protein substrates associated with tumor development (*e.g.* a tumor suppressor p27) is implicated in this process.[86,90] NEDD8 overexpression suppresses those substrates, which can lead to cancer development. Consequently, inhibitors targeting neddylation may be useful for antitumor therapy.[86,87,91]

NEDD8 and ubiquitin both activate E3 ligases. NEDD8 is adenylated by E1 NEDD8-activating enzyme (NAE), then transferred to E1 cysteine side-chain via a thioester linker. Activated NEDD8 is subsequently transferred to the E2 NEDD8-conjugated enzyme called UBC12, then to the CRLs. Conjugation of NEDD8 to these E3 enzyme (the CRLs) triggers ubiquitin ligase activity, causing ubiquitin to be transferred to the ε-amino group of the substrate lysine residues, then polyubiquitination. In general, the process of *ubiquitination* and *neddylation* is similar (Figure 2.2).[86,87,92]

NAE interacts with NEDD8 via three main interfaces (Figure 2.3a). First, NEDD8's acidic face forms salt bridges with the APPBP1 domain of NAE (Figure 2.3b). Second, a hydrophobic patch on NEDD8 contacts UBA3's adenylation domain. Third, the *C*-terminus of NEDD8 penetrates into hydrophobic groove in the UBA3 domain, which closes to the ATP binding site (Figure 2.3c). All three interactions account for ~3350 $Å^2$ contact surface area.[93]

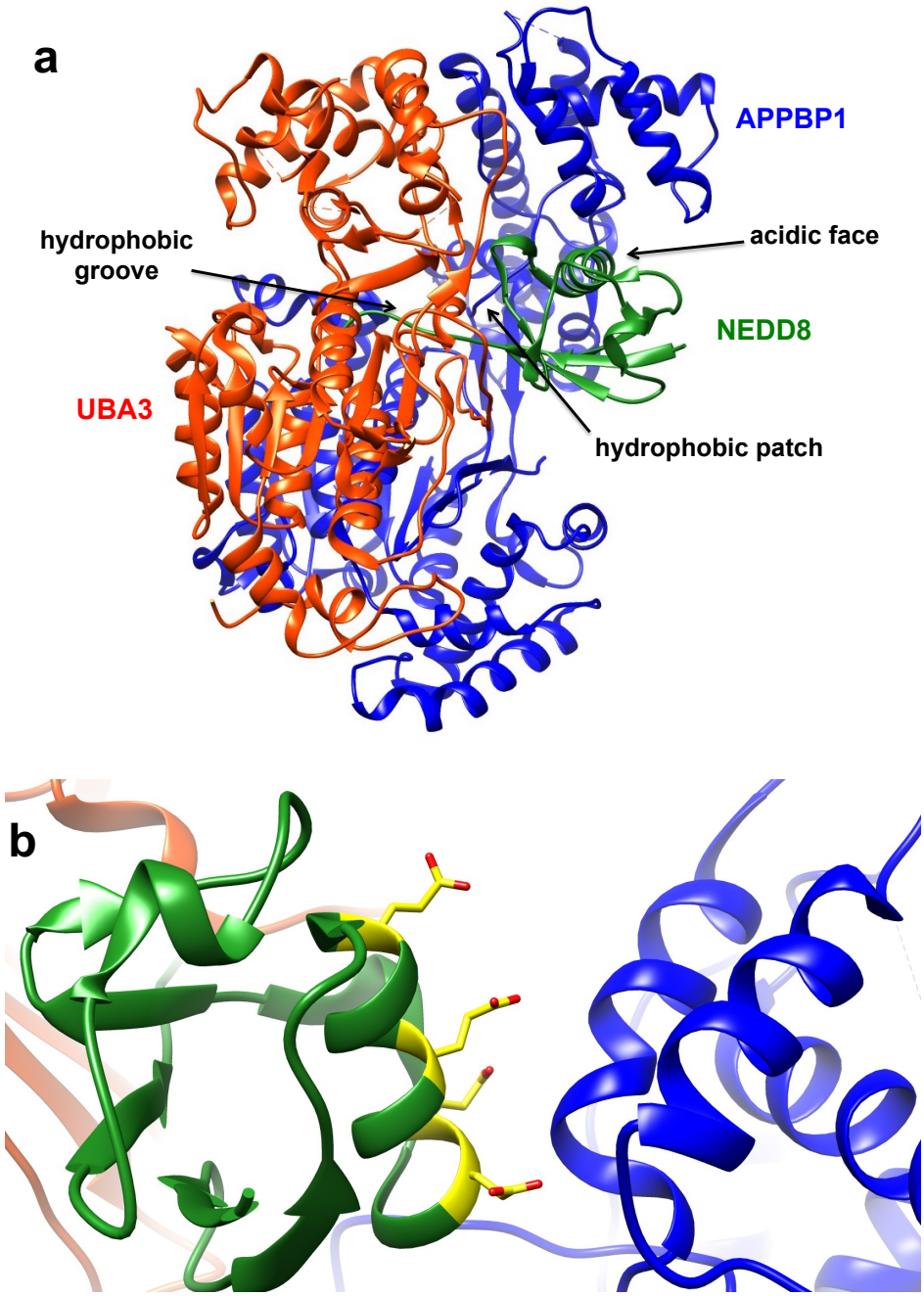


Figure 2.3 (**a**) Crystal structure of NEDD8•NAE (1R4N). (**b**) NEDD8 acidic face interacts with APPBP1. (**c**) Hydrophobic patch and *C*-terminal residues interact with UBA3. Ala72 is highlighted in magenta.

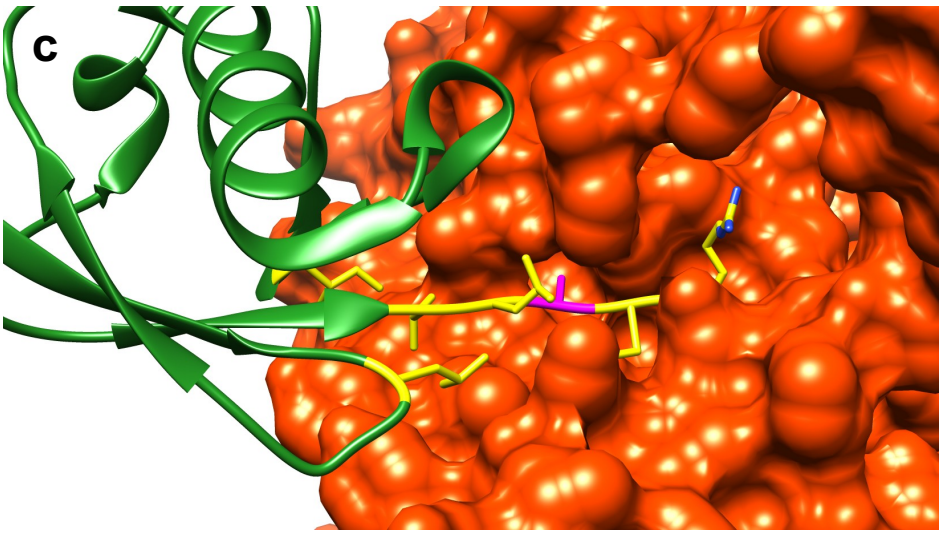
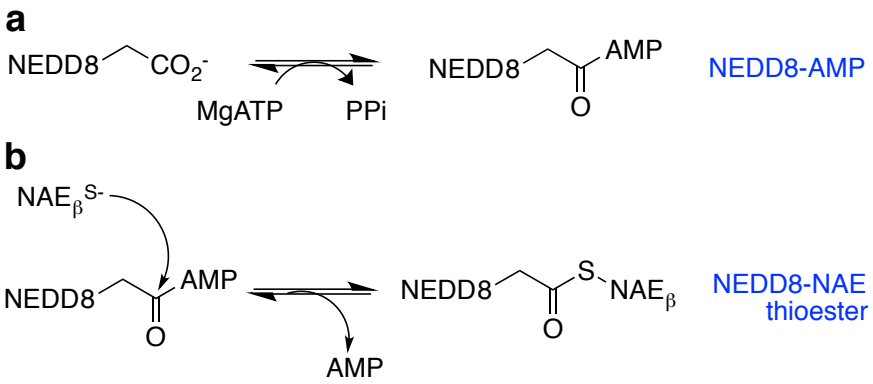
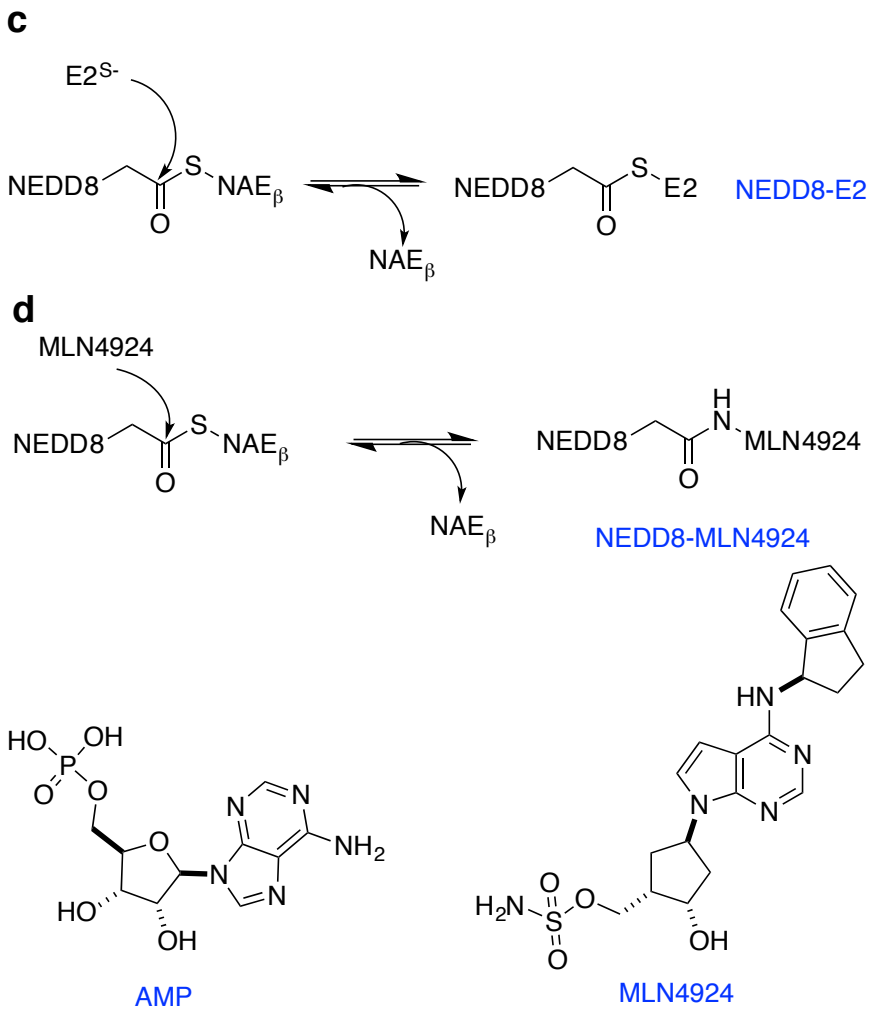


Figure 2.3 Continued.

NEDD8 activation is initiated when its *C*-terminal loop inserts into the UBA3 hydrophobic groove, and the NEDD8 *C*-terminal glycine interacts ATP to form a NEDD8•AMP adduct. The catalytic cysteine of NAE then forms a thioester with NEDD8. Next, a second NEDD8 protein occupies the same hydrophobic groove, which rearranges the NAE structure so that it can transthiolate NEDD8 to E2 enzyme (UBC12) initiating the overall neddylation process (Scheme 2.1a-c).[94-96]



Scheme 2.1 (**a**-**c**) Mechanism of NEDD8 activation by NAE. (**d**) MLN4924 inhibits NEDD8 transthiolation to E2 enzyme.



Scheme 2.1 Continued.

Despite the structural similarity between ubiquitin and NEDD8, they *selectively* interact with their own E1 partners. Specifically, Arg190 of UBA3 acts as a negative gate, which prevents an ubiquitin to bind NEDD8's E1 enzyme due to the steric clash between Arg72 of ubiquitin and Arg190 from UBA3. Conversely, the Ala72 of NEDD8 does not provide additional interaction with Gln190 of ubiquitin's E1 enzyme, hence NEDD8 has reduced affinity for ubiquitin activating enzyme.[97] The selectivity engendered by Ala72 of NEDD8 can be used to design peptidomimetics targeting NEDD8•NAE rather than ubiquitin•E1 enzyme.

MLN4924 (Millennium Pharmaceuticals) is a first-in-class inhibitor of the NEDD8•NAE interaction; it has an IC_{50} of 4.7 ± 1.5 nM.[98] MLN4924, which has a similar structure to AMP, fits into the ATP binding site of UBA3 forming a stable acyl sulfonamide bond with NEDD8 (Scheme 2.1d). The covalent MLN4924•NEDD8 adduct inhibits downstream neddylation, hence deactivates CRLs activities, including protecting the tumor suppressor p27 from degradation. Currently, MLN4924 has passed phase I clinical trials.[99] Apart from this well-known inhibitor, the non-covalent leads **M22** and **ZM223** (from virtual screening) were predicted to bind at the ATP-binding site of NAE similar to MLN4924, and did indeed inhibit the NEDD8•NAE interaction.[100,101]

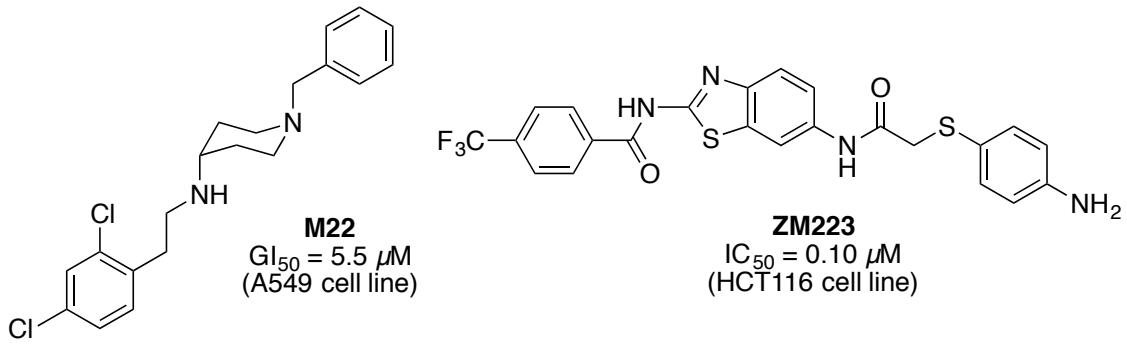


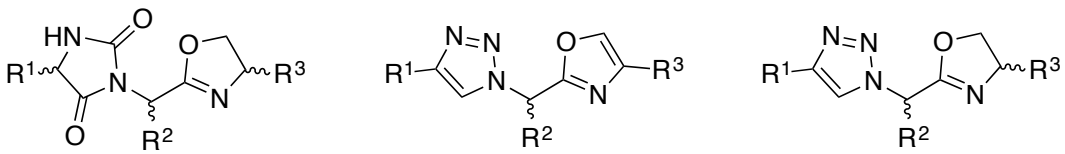
Figure 2.4 Structures of non-covalent compound leads that disrupt NEDD8•NAE interaction.

2.2 Syntheses and Binding of Triazole-oxazole Mimics to Perturb Nef•MHC-I•AP1 Protein-protein Interaction

We analyzed compounds **1** and **2** for overlays on the Nef•MHC-I•AP1 interface (PDBID = 4EMZ)[84] using EKO. Compound **1** showed several hits < 0.35 Å RMSD (Table B.1) and for compound **2** there were two. Compound **1** overlaid at interface regions that appear to be important, including the Nef acidic face, and the sidewall of AP1 that

docked with MHC-I. At this stage, a number of overlays on MHC-I, either between Nef•MHC-I or MHC-I•AP1 interactions, were less than expected. This is because consecutive interfacial side-chains on MHC-I must point to Nef and AP1 alternately (Figures 2.1b-c), but EKO did not consider any overlays of chemotypes that project their side-chains perpendicular or greater to the protein partner. To solve this issue, we used the EKOS routine to evaluate the fit of our chemotypes on the MHC-I peptide *in the absence of the Nef and AP1 proteins*. This approach is not so rigorous, but we hypothesized it might still lead to hits.

Applying EKOS as described above, several overlays of **1** were identified on this MHC-I peptide (Table B.1). Only two overlays for **2** (D-**2vee** and L-**2eev**) had an RMSD <0.35 Å, both were found on the loop at acidic face of Nef (Glu64, Glu65 and Val66; Figure 2.1d). However, syntheses of compounds **1** were highly time-consuming (months to get one compound) hence only compounds **2** were synthesized for this study.



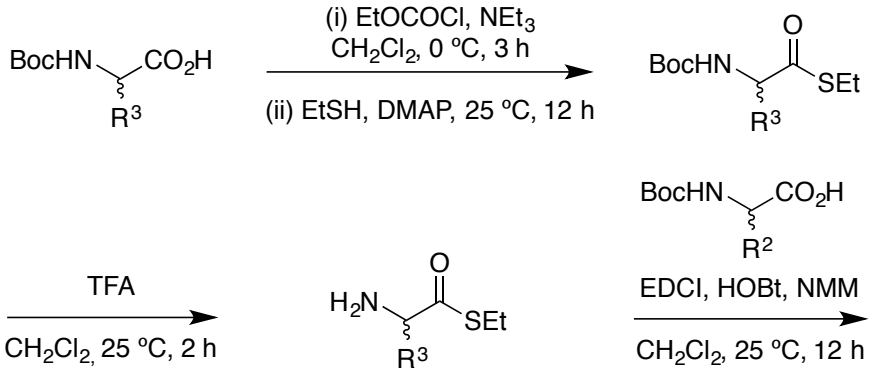
1 **2** **3**

Figure 2.5 Structures of mimic **1**-**3**

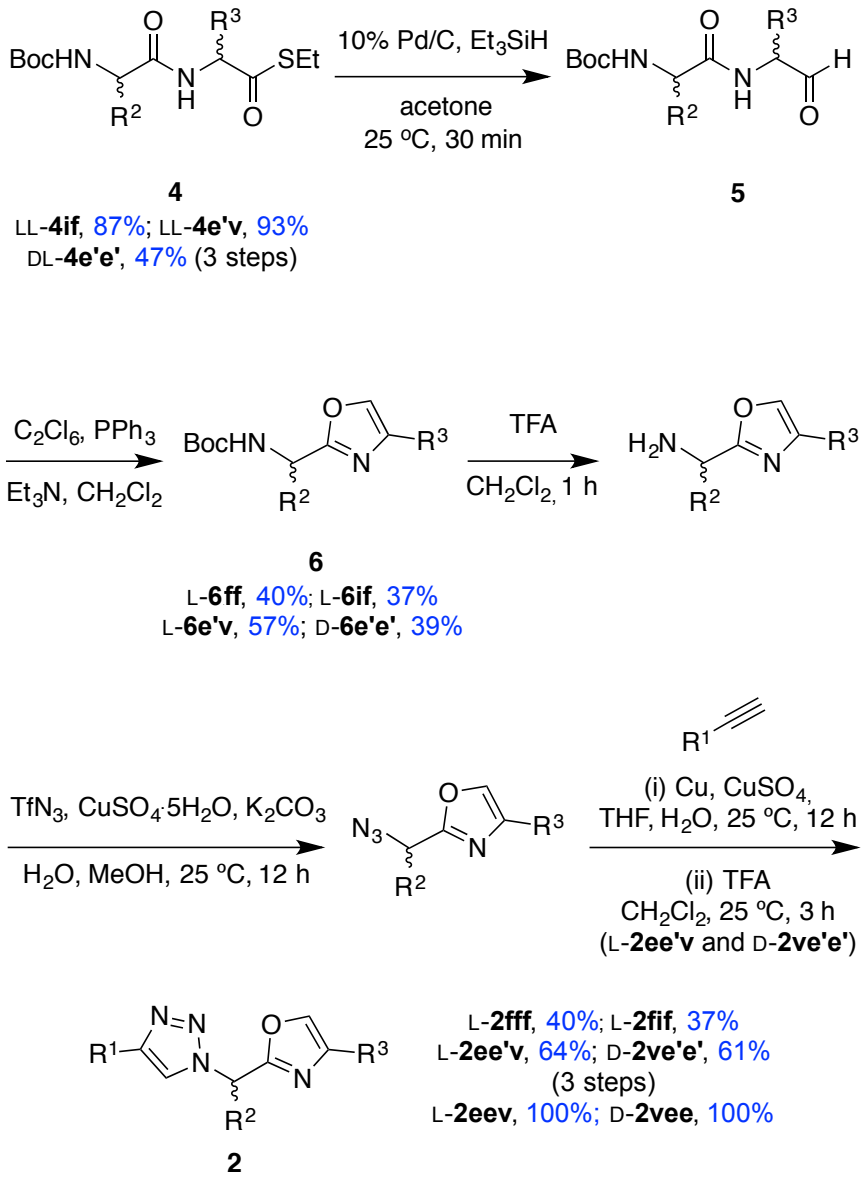
Throughout this work, acid-labile (*tert*-butyl or trityl) side-chain protecting groups were used in the syntheses involving functional side-chain residues. Capital letters (D or L) were used to designate configurations of amino acids. Lower case one-letter codes are used to delineate the amino acid side-chains R^1 – R^3 and relate them to the closest

amino acid; primed letters indicate protected side-chains (*eg* **c'** for the $-CH_2SCH(C_6H_5)_3$ of Cys and **e'** for the $-CH_2CH_2COOC(CH_3)_3$ of Glu).

D-**2vee** and L-**2eev** were prepared from oxazole to triazole as shown in Scheme 2.2. This synthesis involved dipeptide thioester formation **4**, reduction of the -COSEt functionality to aldehydes **5**,[102] then immediate cyclization to oxazole **6**.[103] Finally, *N*-deprotection, conversion to an azide,[104] Cu-mediated triazole formation[105] and side-chain deprotection, gave the desired compounds **2**. Also, the protected forms, D-**2ve'e'** and L-**2ee'v** (di-*tert*-butyl esters) were saved for bioassay. L-**2fff** was synthesized as a "partial negative control", insofar as it contains the same core structure but bears side-chains that were not implicated in the EKO analyses. Another compound, L-**2fif** was also prepared. The purpose of adding this compound was to determine if epimerization at the Ile *C*α during synthesis. In the event, epimerization was not observed (NMR) but the compound was still screened.



Scheme 2.2 Syntheses of triazole-oxazole **2**.



Scheme 2.2 Continued.

Compounds **2** were sent to Dr. Yong Xiong (Department of Molecular Biophysics & Biochemistry, Yale University) to test for disruption of the Nef•MHC-1•AP1 complex (Alphascreen assay; Figure 2.6).[106] Thus, the μ1 domain of AP1 with a His tag was attached to Ni-chelate donor beads, then with the fused MHC-I•Nef complex supported via a FLAG tag, on anti-FLAG acceptor beads. Those two bead types are brought close

together when AP1•MHC-I•Nef forms. The donor beads are excited at 680 nm, which generate singlet oxygen molecules that can only activate the acceptor beads (to emit light at 520-620 nm) when the complex formation draws them into proximity. Donor and acceptor beads without inhibitors give the highest Alphascreen signal, hence this serves as a negative control (no disruption). On the other hand, a blank without any beads (no fluorescent signal at 520-620 nm possible), is a positive control. The inhibitors that can interfere with the Nef•MHC-I•AP1 interaction should provide a fluorescent signal higher than the negative control, but less than positive control.

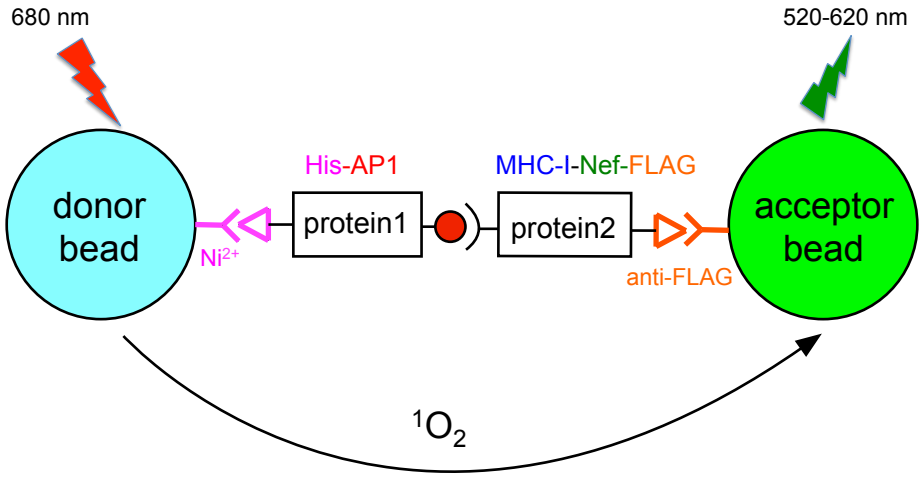


Figure 2.6 Alphascreen assay. Ni-chelated donor bead captures His-AP1 protein, and anti-FLAG acceptor bead captures MHC-I•Nef-FLAG protein. The singlet oxygen molecules generated from donor beads are transferred to the acceptor beads when those come close together.

Six compounds were tested in the Alphascreen assay at 100 μM concentrations. Two of these were from EKO (D-**2vee** and L-**2eev**), two were protected-forms of those hits [D-**2ve'e'** and L-**2ee'v**, e' = -$(CH_2)_2$COOBn], and the two partial negative control (L-**2fff** and

L-**2fif**). Unfortunately, there was insignificant disruption; the best result, L-**2eev**, was about 10% inhibition (Figure 2.7).

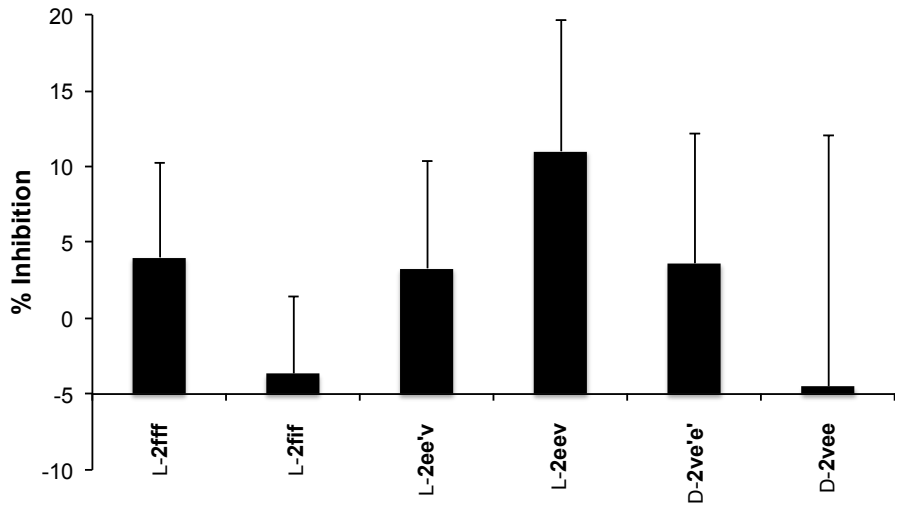
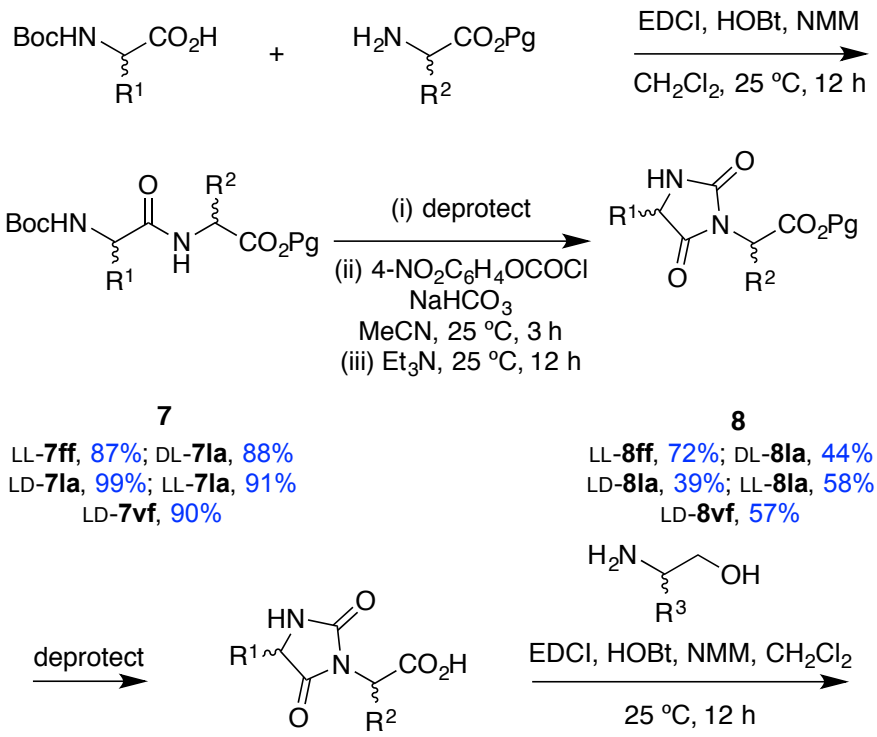


Figure 2.7 Results of an Alphascreen assay. Proteins at 300 nM were premixed with beads (10 μg/mL at final concentration) for 1 h before adding 100 μM inhibitors. Fluorescent emissions were measured using an EnVision 3 plate reader.

2.3 Syntheses and Biological Assays of Hydantoin-oxazoline and Triazole-oxazoline to Perturb The NEDD8•NAE Protein-protein Interaction

NEDD8 inserts its *C*-terminus into the tunnel of UBA3 to initiate neddylation, so our strategy was to occupy that UBA3's hydrophobic groove with chemotpyes. Thus, EKO hits at the *C*-terminus of NEDD8 were selected. Compounds **1** and **3** had three and one promising overlays, respectively (Table B.2), all of which impacted the key Ala72 that governs selectivity over ubiquitin's E1 enzyme. Interestingly, compound **1** gave excellent hits (RMSD ≤ 0.30 Å) around this region.

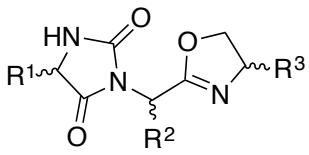
Four chemotypes **1**, including the partial negative control (LLL-**1fff**), were synthesized for analysis as NEDD8•UBA3 inhibitors (Scheme 2.3). Cyclization of dipeptides to hydantoins **8** was based on several literature procedures.[107-109] These hydantoins **8** could be deprotected at the *C*-terminus, then coupled with amino alcohols. Oxazoline cyclization was done in two-steps using methanesulfonyl chloride as an activator. Some other conditions had been tested, but failed to obtain desired products (Figure B.1). Another two compounds, LLD-**1ffs** and LDL-**1vfc'**, were also synthesized, but for other purposes which are not discussed in this chapter.



Scheme 2.3 Synthesis of hydantoin-oxazoline **1**.

9

LLD-**9ffs'**, 68%; LLL-**9fff**, 83%
DLD-**9lal**, 88%; LDL-**9lal**, 63%
LLD-**9lal**, 58%; LDL-**9vfc'**, 85%

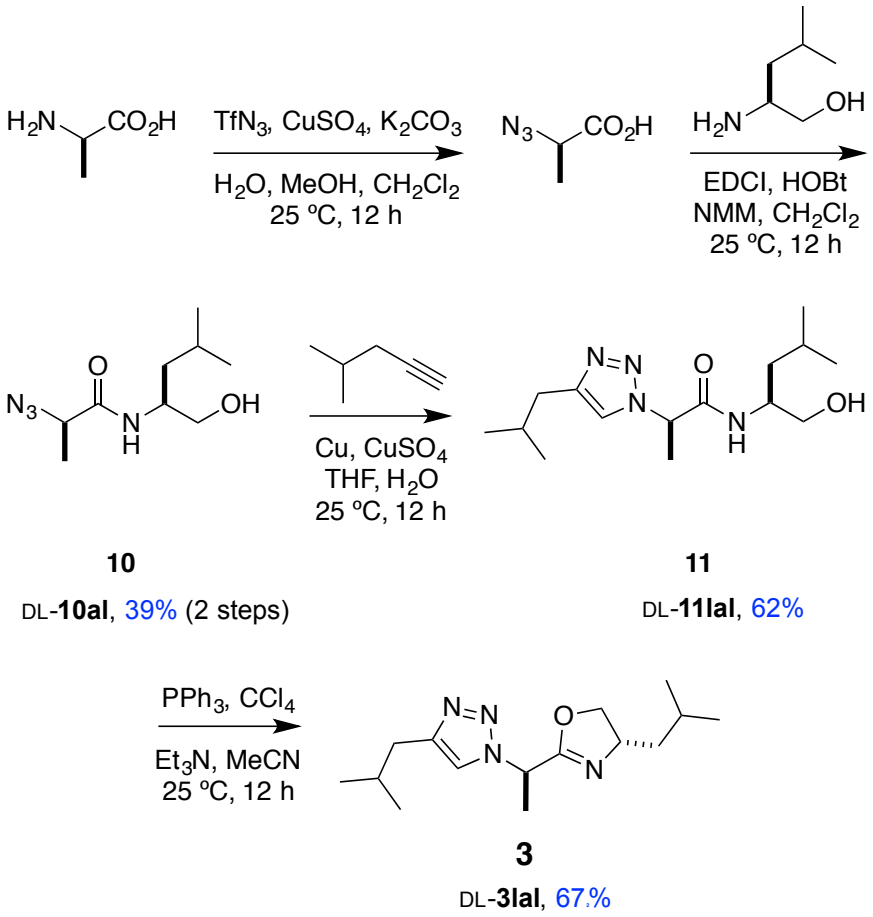


1

LLD-**1ffs'**, 45%; LLL-**1fff**, 39%
DLD-**1lal**, 41%; LDL-**1lal**, 40%
LLD-**1lal**, 39%; LDL-**1vfc'**, 49%
LLD-**1ffs**, 89%

Scheme 2.3 Continued.

One triazole-oxazoline mimic, DL-**3lal**, was synthesized as shown in Scheme 2.4. An azido-dipeptide was prepared by converting the amino to the azido group,[104] then coupling with an amino alcohol. 4-Methyl-1-pentyne was "clicked" to the azido-dipeptide; this particular group was chosen to resemble a leucine side-chain.[105] Finally, alcohol **11** was converted to halide by the Appel reaction,[110] then cyclized to an oxazoline. Overall, synthesis of compounds **3** were easier to achieve than for chemotypes **1** and **2**.



Scheme 2.4 Synthesis of triazole-oxazoline DL-**3lal**

All four positive compounds and one partial negative control were sent to Dr. Jennifer Carew (Hematologic Malignancy Disease Team, The University of Arizona Cancer Center). These compounds were tested on the human AML cell line (MV4-11), which is sensitive to MLN4924 (IC_{50} ~200 nM)[111] by measuring cell viability via an ATPLite assay.[111,112] ATP is a marker for cell viability because it is present in all metabolically active cells but its concentration declines rapidly once at the onset of necrosis or apoptosis. Compounds at 5 different concentrations (up to 10 μM) were incubated with MV4-11 cells for 72 h. Mammalian cell lysis solution was then added to break the cell membrane, and substrate solution (luciferase/luciferin) was added. ATP in the cytosol

drives bioluminescence that can be quantitated in a plate reader. The emitted light is proportional to the ATP concentration: stronger light intensity means more viable cells. Recall that neddylation is a part of tumor suppressor degradation, so compounds that disrupt NEDD8•NAE should reduce cell viability of tumor cells. The results were shown in Figure 2.8. Surprisingly, LLL-**1fff**, a partial negative control, reduced cell viability down to ~ 65% at 10 μM. Other compounds, however, did not show any effect on MV4-11 cell line. This is an unexpected result since LLL-**1fff** does not correspond to any part of the interface implicated in EKO.

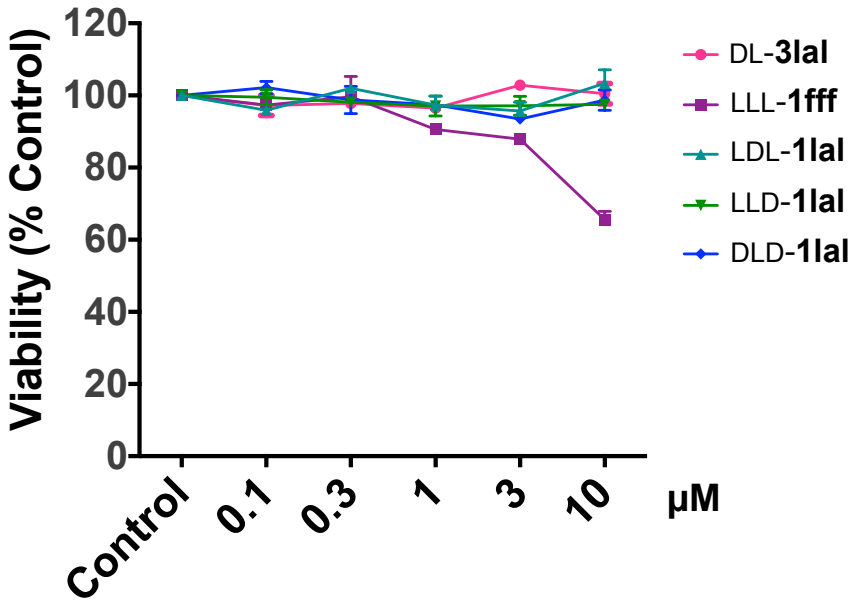


Figure 2.8 Results of cell viability assay. Cells treated at various concentrations of compounds were determined their viability by ATPLite assay.

2.4 Conclusion

EKO algorithm is designed to assist in the design of inhibitors for PPIs by matching chemotypes with interface regions corresponding to the protein ligand. This approach

has been proved to find new inhibitors that disrupt HIV-1 protease dimerization and perturb antithrombin oligomerization[58,59]; nevertheless, this program still needs more examples to build up its reliability. Herein, we chose Nef and NEDD8 proteins to search for further validation. We were attracted to these relatively new targets because of their high value in medicinal chemistry. Based on EKO analyses, compound **1** had conformations that overlaid well on interfaces in the Nef•MHC-I•AP1 complex, and three hits on the NEDD8•NAE interface. However, due to the limited time and synthesis difficulty, only a small number of compounds were synthesized for NEDD8•NAE target. Compounds **2**, which had two hits on Nef's acidic face, were prepared along with two partial negative controls, and compound **3** had one hit for NEDD8 target.

The highest inhibition in the Nef Alphascreen assay, L-**2eev**, was around 10%, which was slightly higher than the protected form. D-**2vee**, on the other hand, had lower inhibition than the protected form. Overall, compounds **2** had little effect on Nef•MHC-I•AP1 interaction. There are possible reasons of these poor inhibitions. First, Nef•MHC-I•AP1 complex has large interface (> 2000 Å), in which only one small molecule (< 900 $Å^2$) may not have enough impact to disrupt protein-protein interaction. Second, MHC-I was fused to *C*-terminal Nef in this assay, which provided a stronger binding interaction than a natural tertiary complex. Lastly, the main interfacial interaction is at MHC-I peptide between Nef and AP1; small molecules mimicking MHC-I should be more interesting because they can intervene between Nef and AP1. With these reasons, the combination of acidic face inhibitors and MHC-I mimics may provide a better inhibition because they cover larger area of PPI interface than only one inhibitor.

In cell viability assay for NEDD8•NAE target, only partial negative control, LLL-**1fff**, reduced cell viability. This is an unexpected outcome since Phe side-chain triads do not

involve in NEDD8•NAE interaction. We hypothesized the cell permeability of compounds was problematic because they must internalize to disrupt NEDD8•NAE interaction. Our collaborator did Western blot assays using cell lysates to test of the compounds could disrupt NEDD8•NAE interaction under these conditions, but they did not (data not shown).

What did we learn from this study? We cannot conclude that EKO did not work in either of these cases, because only a few compounds were prepared, and these were not even the most desirable ones to test. This was so because the compounds were so time-consuming to prepare. However, the Nef•MHC-I•AP1 interfaces are perhaps too large and tightly bound for small molecule disruption. Consequently, this is what we concluded:

(i) Chemotype accessibility is critical. Even if the overlays implicated by EKO at a PPI interface are encouraging, progress will be severely constrained if the compounds cannot be made easily. EKO overlays may not be enough to validate potential mimics precisely since the program only considers $C\alpha$-$C\beta$ alignments, but does not examine binding energies or steric clashes. In the latter project we use Glide to refine our structure designs, but, again, *the analogs must be accessible*.

(ii) Target selection is critical. PPIs with large interface surface areas and featuring tight complexes will be difficult to inhibit.

Despite these limitations, we did not give up on Nef•MHC-I•AP1 or NEDD8•NAE. In fact, the study was stopped because we saw a more interesting opportunity: inhibition of PCSK9•LDLR using more accessible chemotypes, and this is described in the next two chapters.

CHAPTER III

DESIGN OF HYDANTOIN-PIPERAZINE-BASED PEPTIDOMIMETICS TO DISRUPT PCSK9•LDLR INTERACTION

3.1 Introduction

Heart disease, a leading cause of death, is frequently associated with plaque in the arteries causing atherosclerosis, which is attributed by elevated levels of low density lipoprotein (LDL) in the blood. LDLRs (LDL-receptors), on the surface of hepatocytes, are responsible for capturing LDL particles, and importing them into the liver for destruction. Thus, LDLR•LDL complexes on the surface of hepatocytes undergo clathrin-mediated endocytosis into endosomes wherein the acidic environment triggers rearrangement of the LDLR extracellular domain, releasing bound lipoproteins, leaving the LDLR to be recycled to the plasma membrane (Figure 3.1a).[113]

Removal of LDL particles by hepatocytes is negatively modulated by a chaperone called PCSK9[114,115] that recognizes plasma membrane LDLR.[116] PCSK9 inhibits the rearrangement of LDLR in LDLR•LDL complexes, so that *the receptor is not recycled to the plasma membrane but is instead routed to lysosomes where LDL, LDLR and PCSK9 are degraded* (Figure 3.1b).[117-120] Plasma LDL cholesterol levels thus can be *decreased* by inhibiting the PCSK9•LDLR interaction, because that *increases* the display of LDLR on the liver.[113,121,122] This effect was discovered after genetic studies had correlated LDL levels in some patients with mutations associated with gain or loss of function of PCSK9;[123,124] two loss of function mutations correlate with significantly reduced risk of coronary heart disease.[123] Indeed, an individual with no detectable PCSK9, and

extremely low LDL levels, was healthy, suggesting suppression of PCSK9 for lipid lowering is safe.[125] Conversely, individuals with gain-of-function mutations in PCSK9 have a higher risk of coronary heart disease.[123,124]

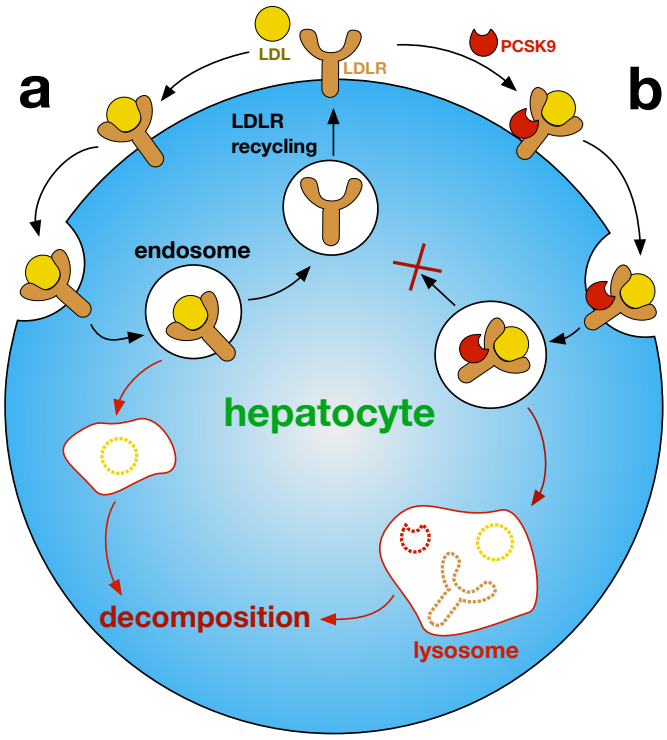


Figure 3.1 (**a**) In the absence of PCSK9, LDL particles are degraded in lysosome, whereas LDLRs are recycled back to the plasma membrane. (**b**) In the presence of PCSK9, all components are degraded in lysosomes.

Multiple clinical studies have shown injectable antibody therapeutics that impede the PCSK9•LDLR protein-protein interaction (PPI) significantly decrease circulating LDL levels.[126] Subsequently, two antibody drugs that disrupt PCSK9•LDLR are FDA approved (Repatha from Amgen and Praluent from Sanofi/Regeneron); they appear to be tolerated well, with no serious side effects, and are efficacious.[127-132] PCSK9•LDLR therefore is a validated target for medicinal chemistry.

The preferred modality for disruption of PCSK9•LDLR, however, is small molecule drugs, not mAbs, on the basis of mode of administration, cost, shelf life, and immunogenic response issues. Thus in 2014 Genentech stated: "Undoubtedly, an orally available small molecule inhibitor of PCSK9, due to lower cost and ease of administration, would be a highly desirable alternative therapeutic agent."[133] Peptides have been used to mimic both components in the PCSK9•LDLR interaction.[134-140] However, peptides are unlikely to provide therapeutic leads for PCSK9•LDLR because they have poor stabilities in plasma.[141] There are few reports of non-peptidic small molecules that inhibit the PCSK9•LDLR interaction. Portola Pharmaceuticals have reported tetrahydroisoquinolines (Figure 3.2a) that increase LDL-uptake in the liver cells, and showed increase of LDLR display on cells.[142] Similarly, Park *et al* reported 2,2-diphenylpropyl derivatives (Figure 3.2b) that increase LDL-uptake, and the decrease of plasma LDL in the wild type mice but not the corresponding PCSK9-knockout mice.[143,144] Nevertheless, both studies did not report any evidence of their compounds disrupting PCSK9•LDLR interaction; those compounds may act via different mechanism. Another small molecule reported by Stucchi et al features an oligo *N*-methyl imidazole (Figure 3.2c) that can disrupt PCSK9•LDLR based on ELISA assay (IC_{50} = 11.2 ± 0.2 μM), and increase LDL uptake in HepG2 cells (EC_{50} = 6.04 μM).[145] Surprisingly, Stucchi's oligoimidazole structure is amphiphilic and featureless, in which it is hard to imagine how this analogue binds on the PCSK9 protein specifically. Thus, in our view, designing small molecules to disrupt PCSK9•LDLR has just begun. Herein, we developed the hydantoin-piperazine **12** (Figure 3.2d) using EKO (as outlined in in chapter I), developed solid-phase syntheses to enable parallel syntheses of analogs, and evaluated their PCSK9•LDLR inhibition using a commercial TR-FRET binding assay kit.

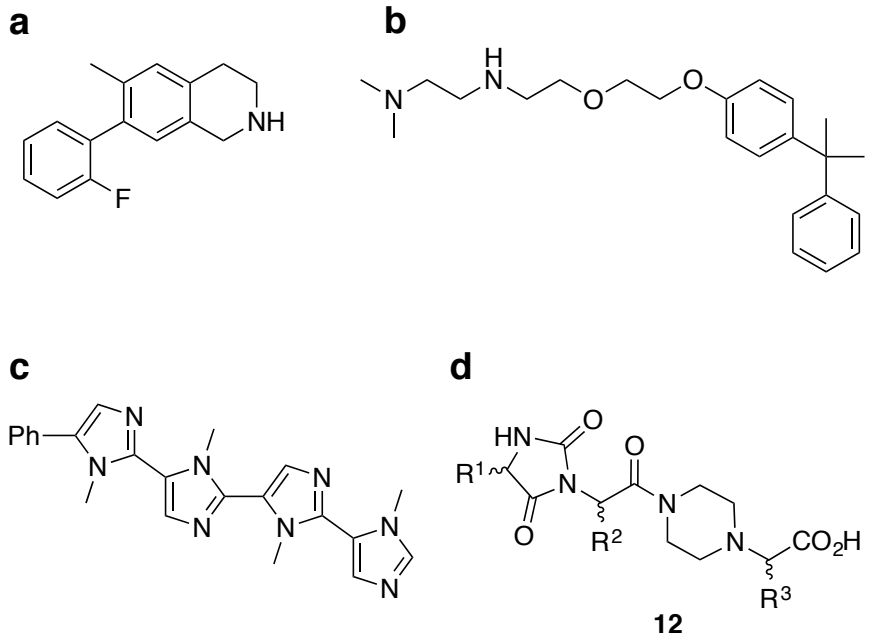


Figure 3.2 Examples of non-peptide small molecules that inhibit PCSK9•LDLR interaction; (**a**) tetrahydroisoquinoline derivative, (**b**) 2,2-diphenyl propane moiety, and (**c**) imidazole-based peptidomimetic. (**d**) Structure of hydantoin-piperazine **12** used in this study.

3.1.1 Proprotein Convertase Subtilisin/kexin Type 9 (PCSK9) Protein: Low Density Lipoprotein Receptor (LDLR)

PCSK9 is a 692-amino acid glycoprotein that is structurally similar to subtilisin proteases.[146-148] It uses its catalytic domain to bind to the extracellular EGF-A domain of LDLR in a 1:1 complex with a K_d, between 170 and 840 nM (Figure 3.3a).[120,149,150]

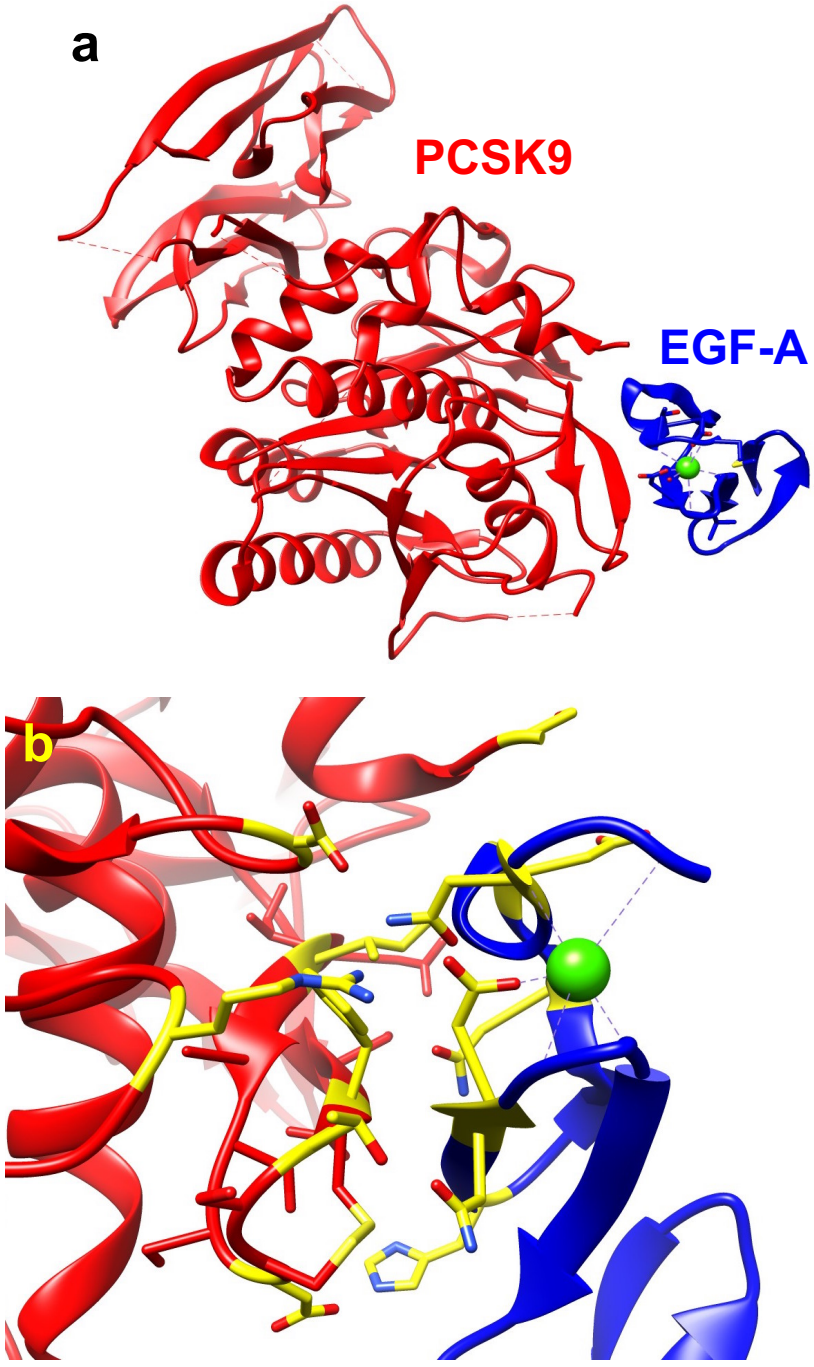


Figure 3.3 (**a**) PCSK9•LDLR(EGF-A) interaction. (**b**) Key interactions are highlighted in yellow.

Residues 367 - 381 and the EGF-A domain of LDLR (residues 293 – 332) interact via a relatively compact interface (*ca* 520 Å^2). The interface features an antiparallel β-sheet

comprised of 377 – 380 of PCSK9 and 308 – 310 of the LDLR EGF-A domain (Figure 3.3b). Additionally, Asp238 and Thr377 of PCSK9 form salt bridges to LDLR EGF-A (to Asn295 and Asn309 respectively), and residues 153 - 155 of PCSK9 also interact with LDLR EGF-A domain. Particularly, Ser153 *H*-bonds to LDLR EGF-A Asp299. Arg194 of PCSK9 forms another salt bridge with Asp310 of EGF-A. In an acidic environment like lysosome, His306 of EGF-A is protonated, which forms an extra salt bridge with Asp374 of PCSK9. Overall, however, the interface is hydrophobic but surrounded by the polar interactions described above. Key apolar amino acids of PCSK9 at this interface include: Ile369, Phe379 and Cys378. Conversely, on LDLR EGF-A Leu318 is critical.[120]

3.1.2 Hydantoins In Minimalist Mimic Design

Imidazolidine-2,4-diones, or hydantoins, are pharmaceutically-interesting scaffolds; the best well-known hydantoin drugs are phenytoin derivatives (Dilantin®, Epilan®, Cerebyx® etc.) which are anticonvulsants. Several hydantoin-based candidates are now in clinical trials for other applications.[151] Several publications and reviews have featured the syntheses of these five-membered heterocycles.[107,151,152] Hydantoin rings have broad structural diversity[151] because they can be modified at the *1*,*3*,*5*-positions; moreover, spirohydantoins and fused hydantoins can be prepared.

Even though hydantoins are synthetically accessible and widely used in pharmaceuticals, there are not many hydantoin-based peptidomimetics. One publication from Hamilton *et al* described preparations of β-strand mimics based on a 1,3-phenyl-linked hydantoin scaffold[153], and another paper from Albericio *et al* featured a hydantoin ring in a short peptide that fit into an active site of caspase-3.[154]

3.1.3 Piperazine-containing Minimalist Mimics

Piperazine is one of the most frequent nitrogen heterocycles in U.S. FDA approved drugs. Most of them contain substituents at both the *N*1 and *N*4-positions.[155] This analogue is significant in the medicinal chemistry of a range of different biological targets.[156] Thus several reports have demonstrated a broad spectrum of pharmaceutical activities of this scaffold, including anticancer, antidepressant, antibacterial and antimalarial agents.[157] This valuable heterocycle has also been incorporated into scaffolds designed to resemble α-helices. König and Maity prepared 1,4-dipiperazino benzene chemotypes to resemble *i*, *i*+3, *i*+7 in a peptidic α-helix, like Hamilton's terphenyl helical mimics.[36] Two publications by Rebek *et al* demonstrated the use of piperazine to prepare compounds in which the side-chains putatively resemble to *i*, *i*+3/*i*+4, *i*+7 residues on α-helix, and the same group prepared other piperazine-containing scaffolds to project side-chains in orientations similar to an extended α-helix (*i*, *i*+4, *i*+8, *i*+11).[158,159]

3.2 Secondary Structure Analysis of Hydantoin-piperazine Mimic 12 by EKOS

Eight stereoisomers of hydantoin-piperazine **12** with all methyl side-chains, *i.e.* those that mimic alanine side-chains (**12aaa**), were evaluated for how similar they project their side-chain orientations when compared to the side-chains orientations of ideal secondary structures. Before doing EKOS analyses, hydantoin-piperazine **12aaa** was subjected to quenched molecular dynamics (QMD)[57] to generate simulated preferred conformations. Briefly, a molecule was minimized and then subjected to molecular dynamic at high temperature (1000K) for a short period time (1500 ps). Molecular

conformations were recorded every 1 ps (i.e. 1500 conformations) then minimized using molecular mechanics.[53,160,161] Conformations within 3.0 kcal/mol of the global minimum were considered as *preferred* conformations and used in the EKOS evaluation. For this mimic **12**, approximately 400-500 conformers of each stereoisomer have energy within 3.0 kcal/mol from the lowest conformation, i.e. those are *preferred* conformers.

Data from the EKOS analysis described above are presented in Figure 3.4a. All eight isomers have representative conformers that match on sheet-type motifs (strand-turn-strand, parallel and anti-parallel β-sheet; all below 0.50 Å) better than helical motifs (3_{10}-, α-, π-helix). The best overlay is a DLL-**12aaa** isomer that spans two parallel β-sheets with an exceptionally low RMSD (0.26 Å, Figure 3.4b). In contrast, mimics **12** do not overlay well on a 3_{10}-helix (Figure 3.4a red bars). Overall, this chemotype has a bias towards sheet-type secondary structure mimicry, although it is not outstanding comparing to other exceptional chemotypes in Chapter V. Although mimic **12** is not outstanding in terms of secondary structure mimicry, it matches well on protein-protein interfaces, e.g. it has 14 hits on PCSK9•LDLR interface. How relevant the secondary structure mimicry of small molecules to the protein-protein interfaces will be discussed in chapter V.

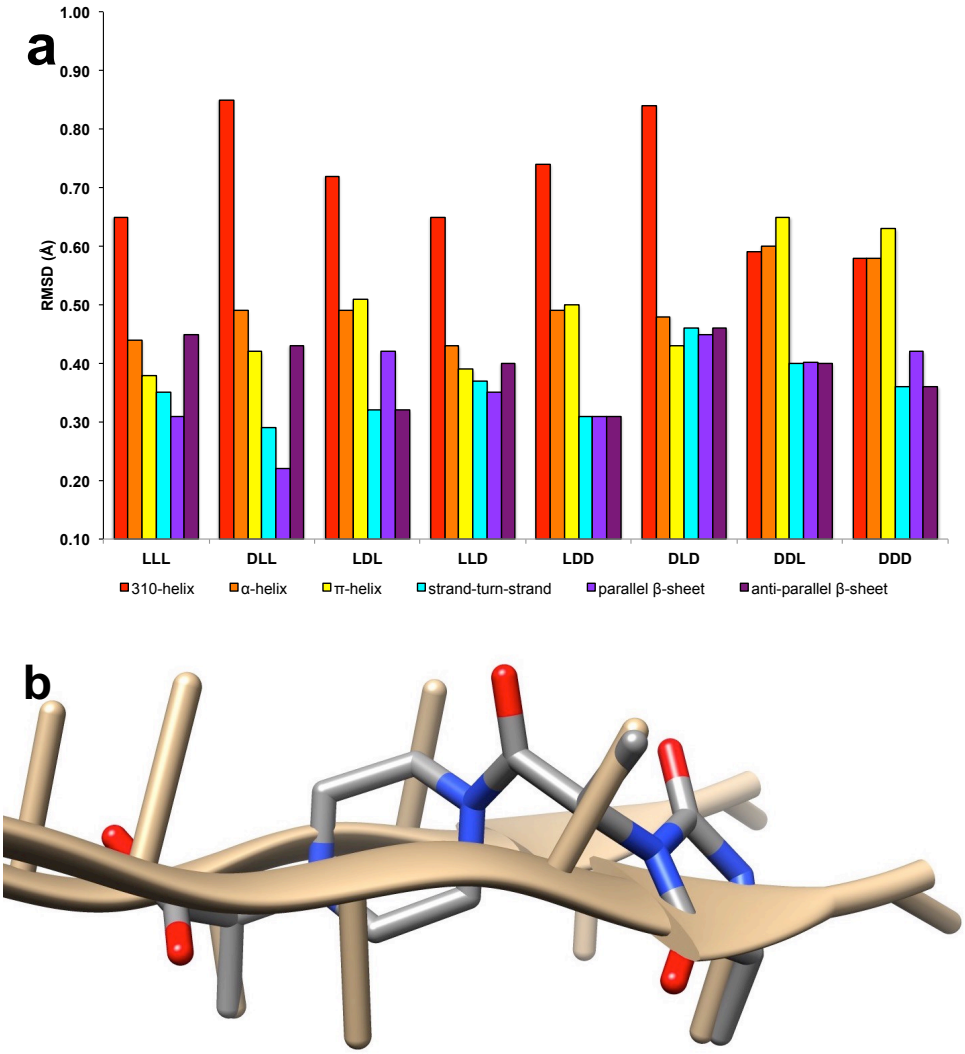


Figure 3.4 (**a**) EKOS matching results of hydantoin-piperazine **12** with ideal secondary structures. (**b**) Overlay of DLL-1**2aaa** on a parallel β-sheet.

3.3 Potential Hits of Hydantoin-piperazine Mimics Derived From EKO Analyses

Simulated preferred conformers generated from QMD were matched on a PCSK9•LDLR crystal structure (PDBID: 3gcx)[162] using EKO.[58] *C*α - *C*β coordinates from the side-chains of hydantoin-piperzaine **12aaa** were systematically overlaid on *C*α - *C*β coordinates for the PCSK9•LDLR interface side-chains. Low root mean square

deviations (RMSDs) between the Cα-Cβ orientations imply preferred conformations of the compound mimics the protein side-chains well. In this work, we set the RMSD cut off at 0.50 Å; overlays below that were considered as "potential hits". A summary of the hits is shown in Table 3.1.

EKO found 12 hits on LDLR, and 2 on PCSK9 shown in Table 3.1. Those hits matched on at least one hot spot residue reported in literature[150], e.g. Asp299, Asn309 and Leu 318 of LDLR, and Ile369 and Phe379 of PCSK9 protein. In this study, we are interested in small molecules targeting PCSK9 protein. This is because small molecules that dock on LDLR may internalize into cells together with LDL particles via LDL-uptake pathway, or may impede LDL•LDLR interaction. To avoid this problem, we discarded two mimics on PCSK9. Of the 12 hits on LDLR, eight hits were chosen for further studies based on availability of starting materials, e.g. mimics involving D-Asn are costly and not suitable to synthesize mimic **12** for the initial screening.

Several crystal structures of PCSK9•LDLR have been reported[149,163,164] but 3gcx was selected based on several criteria: (i) proteins were expressed as the full length (for PCSK9), or enough to maintain their functionality (LDLR); (ii) they are human proteins; (iii) crystal structures were obtained from X-ray diffraction with high resolution (< 3.0 Å); and, (iv) they were crystalized in similar physiological environments (pH 7.4). Another crystal structure, 3bps,[150] was obtained under almost the same conditions, both structures are nearly identical (RMSD = 0.26 Å for all atoms), and EKO analyses showed the results from those two crystal structures are almost identical.

Table 3.1 Potential LDLR mimics based on scaffold **12**, from EKO analyses RMSD ≤ 0.50 Å.

Match on	Isomer	Residues	RMSD (Å)
LDLR	DDD	D299-L298-N301	0.39
	LDD	D299-L298-N301	0.41
	LLD	C297-N301-D299	0.42*
	DDL	D299-L298-N301	0.43
	LLL	C297-N301-D299	0.44*
	LDL	D299-L298-N301	0.45*
	DLD	C297-N301-D299	0.46*
	LLL	L298-D299-N301	0.46*
	DLL	C297-N301-D299	0.47*
	DDD	V307-C308-L318	0.47*
	LLD	L298-D299-N301	0.48
	DLL	N309-C308-L318	0.50*
PCSK9	DLL	F379-V380-I369	0.48
	DLD	F379-V380-I369	0.48

* Selected hits

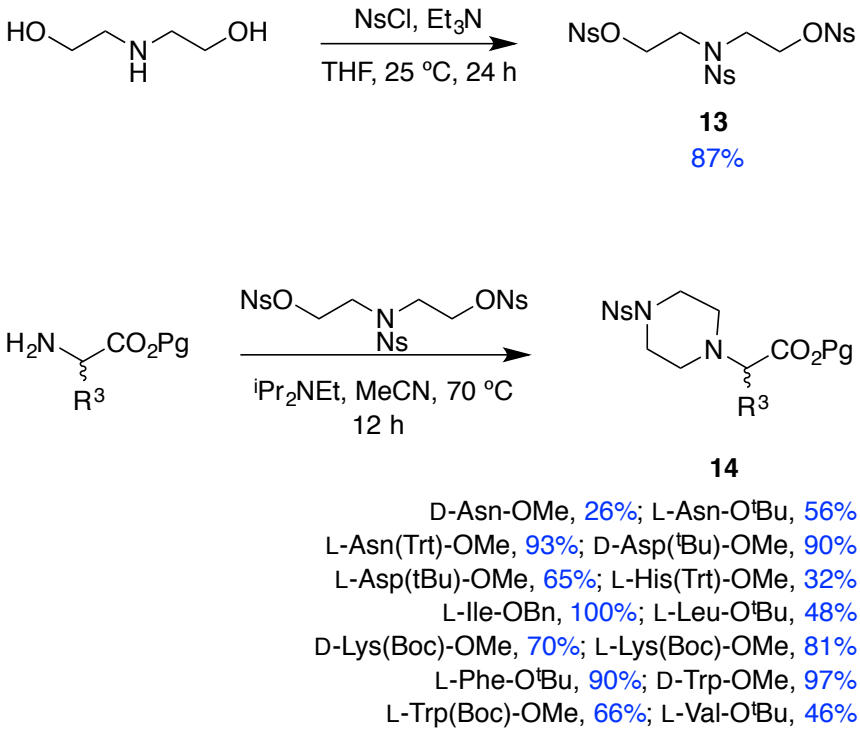
3.4 Syntheses of Piperazine-based Amino Acids 'Pip-acids'

Piperazine-amino acids 'pip-acid' **15** is a relatively new building block that is hardly represented in the literature,[165] but in Dr. Burgess's group is researching syntheses and applications of this fragment. This unit has some interesting features. A piperazine ring imparts rigidity relative to oligopeptides. A wide range of pip-acids can be synthesized from both natural and non-natural amino acid precursors, hence there are opportunities to prepare libraries of pip-acids for drug discovery, e.g. for fragment-based screening. Additionally, the piperazine ring is common in pharmaceuticals as described

above.[156,166,167] Thus piperazine-amino acids are valuable for synthesizing hydantoin-piperazine molecules (my project), but also as valuable building blocks for other applications (Maritess Arancillo in Dr Burgess' group).

Pip-acids **15** in Scheme 3.1 were prepared from *C*-protected amino acids and trinosyl diethanolamine. Trinosyl diethnolamine **13** was prepared from commercial diethanolamine and 4-nosyl chloride under basic conditions.[168] We successfully made this material on ~100 g scale, and purified it by precipitation in dichloromethane and further recrystallization in THF/MeOH mixture. Purified trinosyl diethanolamine **13** was reacted with *C*-protected amino acids in refluxing acetonitrile to give the protected piperazine-amino acids **14**. We modified a protocol that similarly used aryl amines in the literature, so that we could synthesize pip-acids without microwave irradiation.[169] Several different protecting groups on various amino acids are tolerated under the conditions we developed. In the final step, protecting groups at the main-chain carboxylic acid were removed by different methods to obtain acid-free pip-acids **15** as the building blocks for solid-phase syntheses. Overall, it is possible to prepare this analogue with any kind of amino acids; even the sensitive side-chain of histidine could be synthesized by this protocol. Nevertheless, there are some cases that reactions do not proceed smoothly. For instance, D-Asn-OMe is prone to methylate at amide side-chain during syntheses, which can be avoided by using a side-chain protecting asparagine, e.g. trityl protecting group; however, due to the cost of D-amino acid, we did not prepare further. Another amino acid, L-Trp(Boc)-OMe, lost its Boc group during hydrolysis; however, this indole-deprotected product is useful in solid-phase syntheses since an indole is less reactive than a primary or secondary amine with respect to amide bond formation reactions. If necessary, Boc-protecting groups can be re-introduced onto the indole ring after hydrolysis. Benzyl-protecting groups cannot be cleaved via

hydrogenolysis in syntheses of pip-acids because that would reduce the nosyl-nitro group. Consequently, we used BBr_3 to remove benzyl groups without affecting nosyl group or the *C*α chirality.

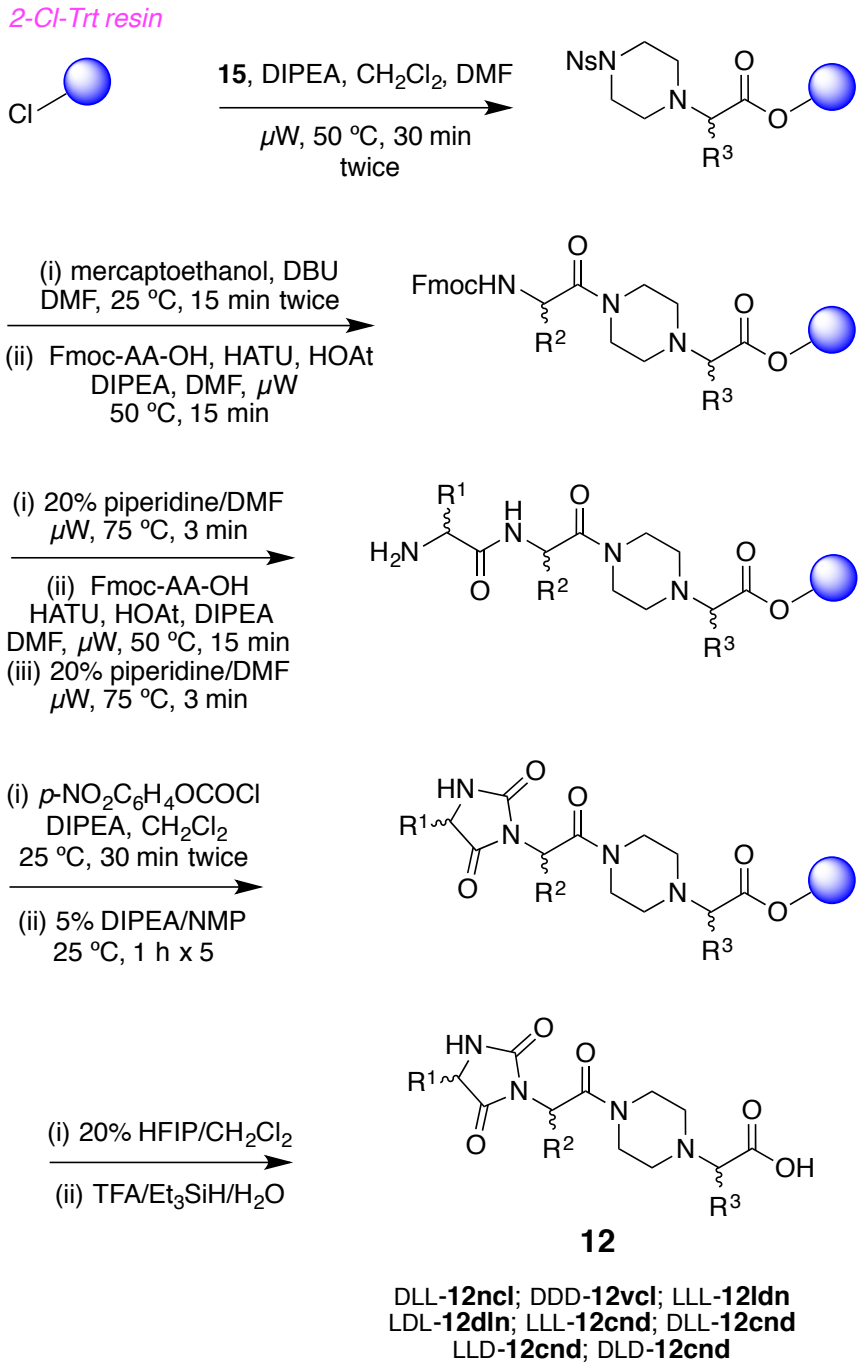


Scheme 3.1 Syntheses of piperazine-amino acids 'pip-acids'.

3.5 Syntheses of Hydantoin-piperazine Mimics 12 via Solid-phase Syntheses and TR-FRET Binding Assay

Throughout this work, acid-labile (*tert*-butyl or trityl) side-chain protecting groups of Fmoc-amino acids were carried through the pip-acid syntheses unless otherwise indicated. Similar to previous chapter, capital letters (D, L) are represented configurations of side-chain residues, and lower case one-letter codes are presented the amino acid side-chains R^1 – R^3 with primed letters indicate protected side-chains (*eg* **c'** for the $–CH_2SCH(C_6H_5)_3$ of Cys and **n'** for the $–CH_2CONHCH(C_6H_5)_3$ of Asn).

Hydantoin-piperazine mimics **12** were synthesized via solid-phase syntheses (Scheme 3.2) on 2-chlorotrityl resin to produce *C*-terminal carboxylates after cleavage. Loading the first pip-acid was achieved under basic solutions using microwave irradiation to accelerate the reaction. Nosyl group was deprotected using mercaptoethanol and DBU, and then two cycles of standard coupling/deprotection of Fmoc amino acids were performed under microwave irradiation. After the final Fmoc deprotection, the amino group was reacted with *p*-nitrophenyl chloroformate, then the adduct was treated with Hunnig's base to initiate cyclization to hydantoins. Finally, the compounds were cleaved from resins under mildly acidic conditions to preserve the side-chain protecting groups; this facilitated detection during the preparative HPLC relative to the core structure that does not contain chromophores. Purified compounds were deprotected from their acid-labile protecting groups in a $TFA/Et_3SiH/H_2O$ cocktail, precipitated, washed several times with cold ether to remove byproducts, then lyophilized to obtain the final products as powders. Two compounds, LLL-**12ldn** and LDL-**12dln**, were synthesized using benzyl-protecting aspartate ($-CH_2COOBn$), which were removed via hydrogenolysis with 10% Pd on charcoal under H_2 atmosphere.



Scheme 3.2 Solid-phase syntheses of hydantoin-piperazine peptidomimetic **12**.

All eight compounds were evaluated for their PCSK9•LDLR inhibitions using a TR-FRET assay kit (BPS Bioscience). In this assay, the EGF-AB domain of LDLR is labeled with europium (Eu^{3+}) cryptate, which has an excitation/emission wavelengths at 320 and 620

nm respectively. This lanthanide complex has a fluorescent decay longer than background fluorescence or light scattering from media, thus the fluorescent intensity can be measured after interfering signals have completely decayed. This LDLR-Eu^{3+} interacts with the protein partner, PCSK9, which has been biotinylated. Another component, an acceptor conjugated with dye, recognizes the biotin on PCSK9 protein. This dye, usually an allophycocyanin derivative, can be excited around 620 nm and emit fluorescence at 665 nm. When LDLR and PCSK9 interact together (i.e. a positive condition), the europium complex and the acceptor dye come into close proximity, resulting the Förster Resonance Energy Transfer (FRET) that gives a fluorescence maxima at 665 nm. If the two proteins do not interact (i.e. a positive outcome if caused by a small molecule), the fluorescent intensity of europium is greater at a shorter wavelength (620 nm).

In this experiment, we set the sample containing LDLR, PCSK9, dye acceptor and vehicle (DMSO) as the negative giving a highest intensity of 665 nm wavelength. For the positive, PCSK9 was excluded, which reduced the intensity at 665 nm but increased fluorescence intensity at 620 nm. Samples containing PCSK9•LDLR inhibitors should show a higher intensity at 620 nm than the negative control, but not exceed the positive one. Pep2-8, a 13-residue peptide developed by Genentech,[133] was used here as a positive control. We screened all eight inhibitors at 10 μM and 50 μM concentrations (Figure 3.5). The results showed three mimics could inhibit PCSK9•LDLR above 20 % at 50 μM comparing to negative control (DDD-**12vcl**, LLL-**12ldn**, LDL-**12dln**); however, the deviations were huge, and some tests had irregular inhibition outcome. For instance, DLL-**12ncl**, DDD-**12vcl**, DLL-**12cnd** and DLD-**12cnd** had fluorescent intensity lower than negative control (negative % inhibition). Moreover, inhibitions of DLL-**12cnd** and LLD-**12cnd** at 10 μM were slightly better than 50 μM. All these events raised a

question about the quality of this assay. Nevertheless, the data implied that our small molecules might have poor binding affinities toward PCSK9 protein, or that the TR-FRET assay had insufficient sensitivity to make any conclusions. In either case, modification of the chemotype is necessary to improve inhibitory potency, and other assays were developed, as described in the next chapter.

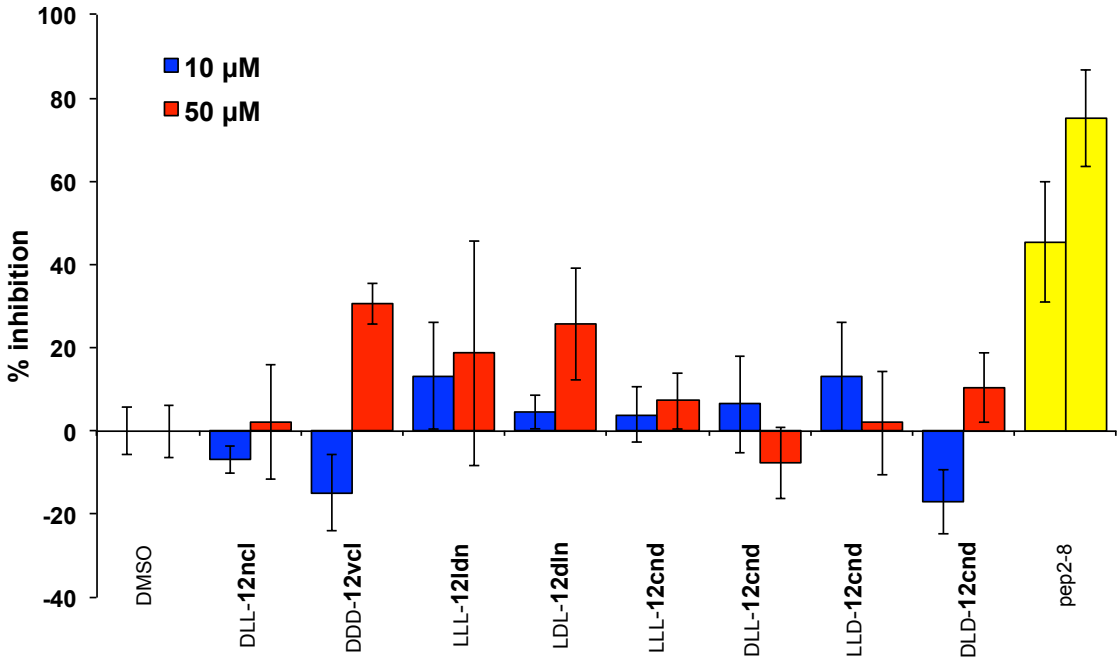


Figure 3.5 Compounds **12** were screened for their inhibition using a commercial PCSK9•LDLR TR-FRET assay kit (BPS Bioscience). Compounds **12** and Pep2-8 were tested at 10 and 50 µM.

3.6 Conclusion

Diseases related to high levels of bad cholesterol, cLDL, are major human threats. Traditional drugs lowering total cholesterol in plasma, e.g. statins, are not effective for

every patient. One reason is because statins can promote the excretion of PCSK9 that reduces LDL receptors on liver cell surface.[170,171] In turn, LDL particles cannot be cleared out from the plasma, leading to the plaque formation and subsequently causing atherosclerosis. New drugs targeting PCSK9 might lower LDL plasma concentrations, and might be used alone or in combination with statins. Recently, Repatha and Praluent became the first approved mAb drugs targeting PCSK9. However, the cost and mode of administration of these therapeutics have encouraged researchers to develop small molecules that impede PCSK9 functionality.

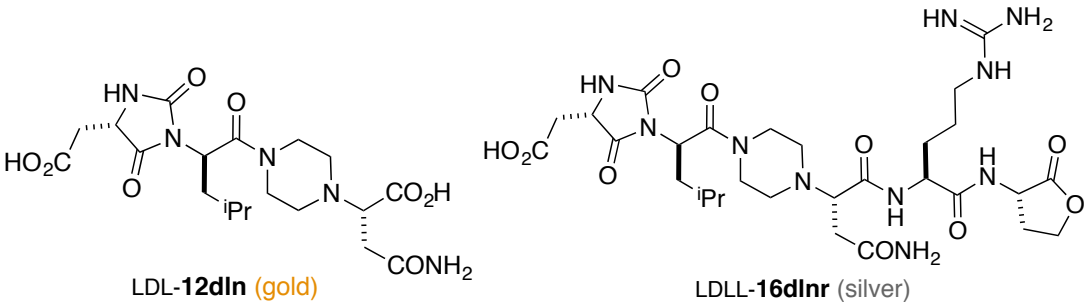
Hydantoin-piperazine mimic **12** were designed as a novel small molecule to disrupt PCSK9•LDLR. Both hydantoin and piperazine-amino acid give rigidity to **12**, and are both common substructures in pharmaceutical drugs. Analysis by EKO revealed many hits for the PCSK9•LDLR interaction. To validate the EKO results, we synthesized pip-acids **15**, and used them to construct hydantoin-piperazine mimics by solid-phase syntheses; compounds bearing the side-chains implicated by EKO analyses were prepared. All 8 mimics were screened together with the positive peptide, Pep2-8, using a TR-FRET assay kit, but the results were inconclusive; the best three compounds, DDD-**12vcl**, LLL-**12ldn** and LDL-**12dln**, showed % inhibition greater than 20%, but the deviations were huge, even Pep2-8 data showed nearly 20% deviation. Surprisingly, fluorescent intensities of some cases were even lower than the negative control. It is not clear whether these peculiar results are due to the poor activities of mimics, or the assay itself is not suitable for small molecules that may have weak binding affinities. In any event, we decided to modify mimics **12** to improve their binding affinity toward PCSK9 by using Glide to refine our structures. This process led to another set of new potential hits that, after a series of experiments, show lead compounds to disrupt PCSK9•LDLR interaction. These data are described in the next chapter.

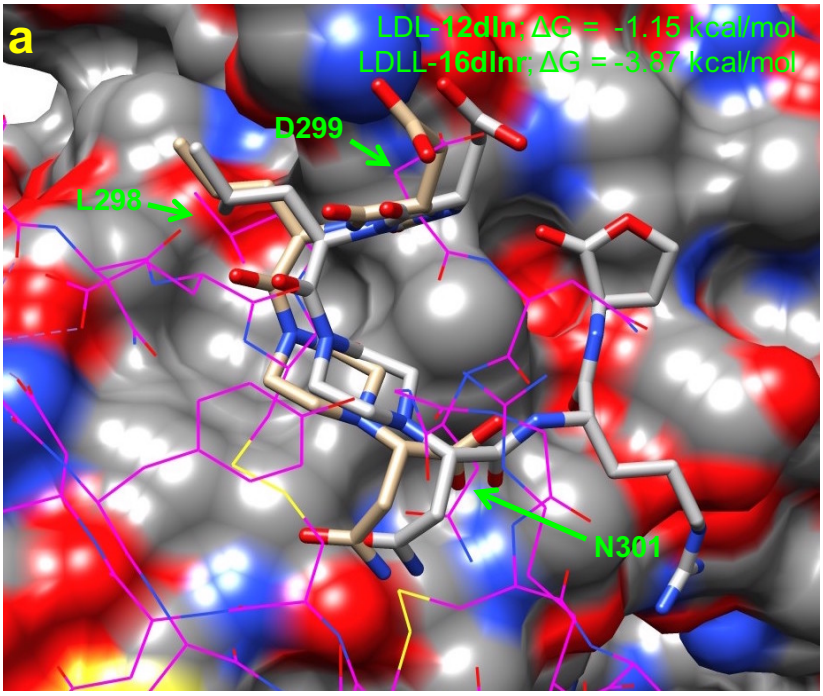
CHAPTER IV

OPTIMIZATION OF HYDANTOIN-PIPERAZINE-BASED MIMICS TO IMPROVE INHIBITORY EFFICACY AGAINST PCSK9•LDLR INTERACTION

4.1 Introduction

In the previous chapter, compounds **12** did not show significant inhibitory activities against PCSK9•LDLR interaction in the TR-FRET assay. We hypothesized compounds **12** might have weak binding affinities, so our next strategy was to do structural modifications. Consequently, an iterative docking and energy minimization procedure was used to virtually modify chemotypes **12** to include additional pharmacophores that would increase binding affinities. Specifically, the preferred conformations of **12** were overlaid on LDLR in the 3gcx structure, LDLR was removed, then the chemotypes were docked in place using Glide within the Schrodinger package.[172-174] This procedure gave "baseline energies" for interactions of **12** with PCSK9 to which data from docking virtually modified analogs could be compared;[175] this procedure was performed in the Schrodinger software package using the CombiGlide routine. Figure 4.1 shows an example of docked and minimized structure of **12**•PCSK9 generated from the original EKO overlay (gold color). We noticed that there is a conspicuous negatively charged cavity in PSCK9, proximal to the *C*-terminus of **12**, but not interacting with it. Consequently, we made that cavity a priority in CombiGlide simulations. Those calculations indicated the optimal Arg at the **12** *C*-terminus fits into that negative cliff on PCSK9 leading to compound **16** (silver color), which improved calculated binding energy by 2.72 kcal/mol due to the extra H-bonds to Asp^{367}.





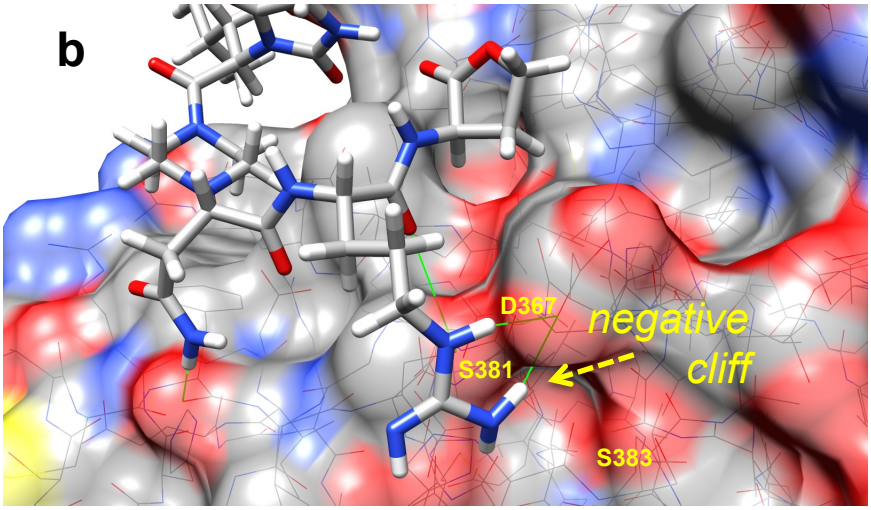


Figure 4.1 Docking of LDL-**12dln** (gold) and LDLL-**16dlnr** (silver) onto PCSK9. (**a**) Compounds LDL-**12dln** and LDLL-**16dlnr** overlay on the LDLR (shown in magenta wire),

in which six $C\alpha$-$C\beta$ atoms of chemotypes and LDLR side-chains (green arrows) are compared; RMSD = 1.75 and 1.95 Å for LDL-**12dln** and LDLL-**16dlnr**, respectively. (**b**) Improving binding affinity of LDLL-**16dlnr** is likely due to the H-bonds of Arg residue to the Asp^{367} of PCSK9 at a negative "cliff face" region as indicated in green lines.

We refined our eight *first-generation* hydantoin-piperazine mimics **12** with CombiGlide to obtain fifteen *second-generation* mimics **16** shown in Scheme 4.1 and Scheme 4.2. These compounds were tested both in binding and cellular assays, and the results indicated the promising lead, LDLL-**16dlnr**, that had dissociation constant (K_d) at 25 μM based on SPR analysis, and showed significant increase of fluorescent-labeled LDL uptake in a LDL-uptake assay.

4.1.1 Glide and Combiglide

Methods for docking flexible ligands on the rigid 3D receptors, especially from high-resolution crystal structures, are extensively employed in drug discovery. Numerous programs are available for academic or commercial purposes, in which they have similar steps of docking ligands on the receptors. Programs search the lowest-energy conformations of a ligand, and then calculate their binding interactions to the protein receptor. Those interactions are the sum of major binding energies, e.g. van der Waals, Coulombic, hydrogen bonds and electrostatic energies. Those calculated energies are processed through docking score algorithms to rank how good each chemotype docks on protein receptors. However, solvent effects, entropic effects and side-chain flexibility of protein receptors can perturb the docking calculations leading to inaccuracy of results. Some programs have been calibrated with experiments to minimize those issues, e.g. MOE-Dock, GOLD, Glide, FlexX etc.[176]

Glide[172,173] (Grid-based Ligand Docking with Energetics) has been designed to perform an exhaustive search of the positional, orientational, and conformational space available to the ligand. It produces a set of initial ligand conformations to roughly screen on the receptor to search for promising ligand poses. Selected promising pre-screen poses are minimized in the field of receptor using the molecular mechanics energy function, OPLS-AA (Optimized Potentials for Liquid Simulations-All Atoms) force field.[175] Lowest-energy poses are subjected to a Monte Carlo calculation to further minimize torsional strains. Finally, the scoring function, GlideScore, calculates binding affinity and reports a docking score, which can be used to compare among other compounds in database screens. GlideScore has two standard algorithms; a softer, more forgiving, 'SP' calculation that minimizes false negatives suitable for a huge database, and a harder 'XP' function that minimizes false positives for a limited number of optimized compounds.[174]

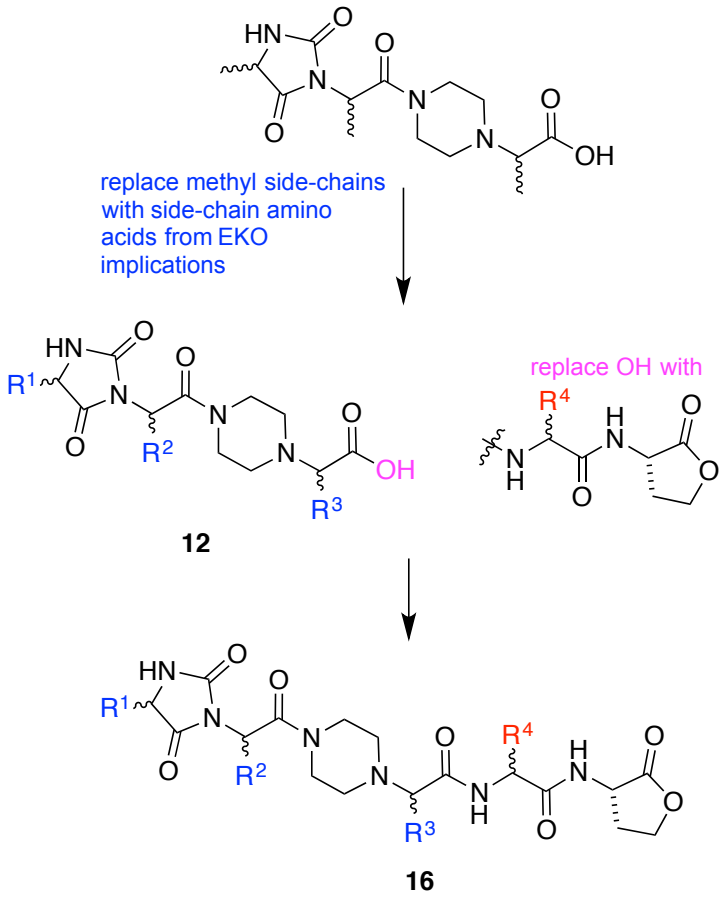
Virtual combinatorial ligands can be enumerated from the parental ligand using CombiGlide. Fragment libraries are replaced at certain positions on parental ligands to generate the virtual combinatorial libraries (for example see Scheme 4.1), and then those libraries are docked on protein receptors with Glide to rank the fragments that should be incorporated. For a large virtual combinatorial library, e.g. > 300 enumerated ligands, the fast SP calculation is recommended to rank those enumerated ligands while minimizing computational cost, and then the small subset (~ top 10-30 ligands from the SP ranking) is recalculated using the precise XP calculation to obtain high-quality ranking results. This process is useful to deduce a number of potential ligands to be synthesized for study.

4.2 Design Second Generation Hydantoin-piperazine Mimics with Glide

Molecular dockings of the virtual libraries were performed using Glide in Schrödinger package (version 2015-4).[172-175] All eight selected hits from the first-generation hydantoin-piperazine were used as the templates for optimizations. First, methyl side-chains of the hydantoin-piperazine templates were substituted with the side-chain amino acids from EKO matchings (Table 3.1) corresponding to the side-chain residues at LDLR interface. Docking of these on PCSK9 would give the baseline energies for the interactions between **12** and PCSK9. Modification at *C*-terminus was done using CombiGlide routine. In this work, we substituted carboxylic acid with amino-acid lactones (Scheme 4.1). We made our own library by varying R^4 with both *L*- and *D*-genetically-encoding amino acids (except Met, Gly and Pro) and an additional ornithine amino acid for the total of 36 virtual fragments. These virtually modified compounds **16** were docked on PCSK9 again. We used OPLS 2005 force field with 20 Å of the grid box, and restricted the virtual conformations within 5 Å from the parental poses to maintain the original poses obtained from QMD minimization. We picked the mimics that had best docking scores from each template for further study. In most cases, the positive-charge residues, His, Lys or optimally Arg, were the best for R^4 position, and binding poses were comparable to the original poses from QMD.

Based on our experience preparing compounds **12**, cyclization of hydantoin was difficult when side-chain R^1 was cysteine, and the final product had relatively poor stability. We decided to replace this cysteine residue by screening other amino acid side-chains using CombiGlide and Glide routines. The top results showed glutamine or glutamate derivatives gave the best docking score (highest negative Gibbs free energy).

Therefore, we substituted Cys on hydantoin with those Glu or Gln (for instance, from DLL-**12cnd** to DLLL-**16qndr**).



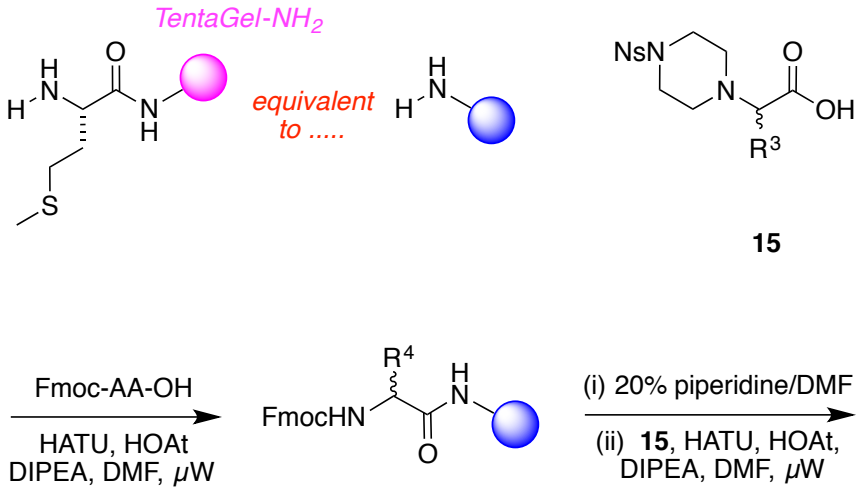
Scheme 4.1 Illustrative process on how to modify mimic **12** to **16**.

4.3 Syntheses of Second Generation Hydantoin-piperazine Mimics via Solid-phase Syntheses

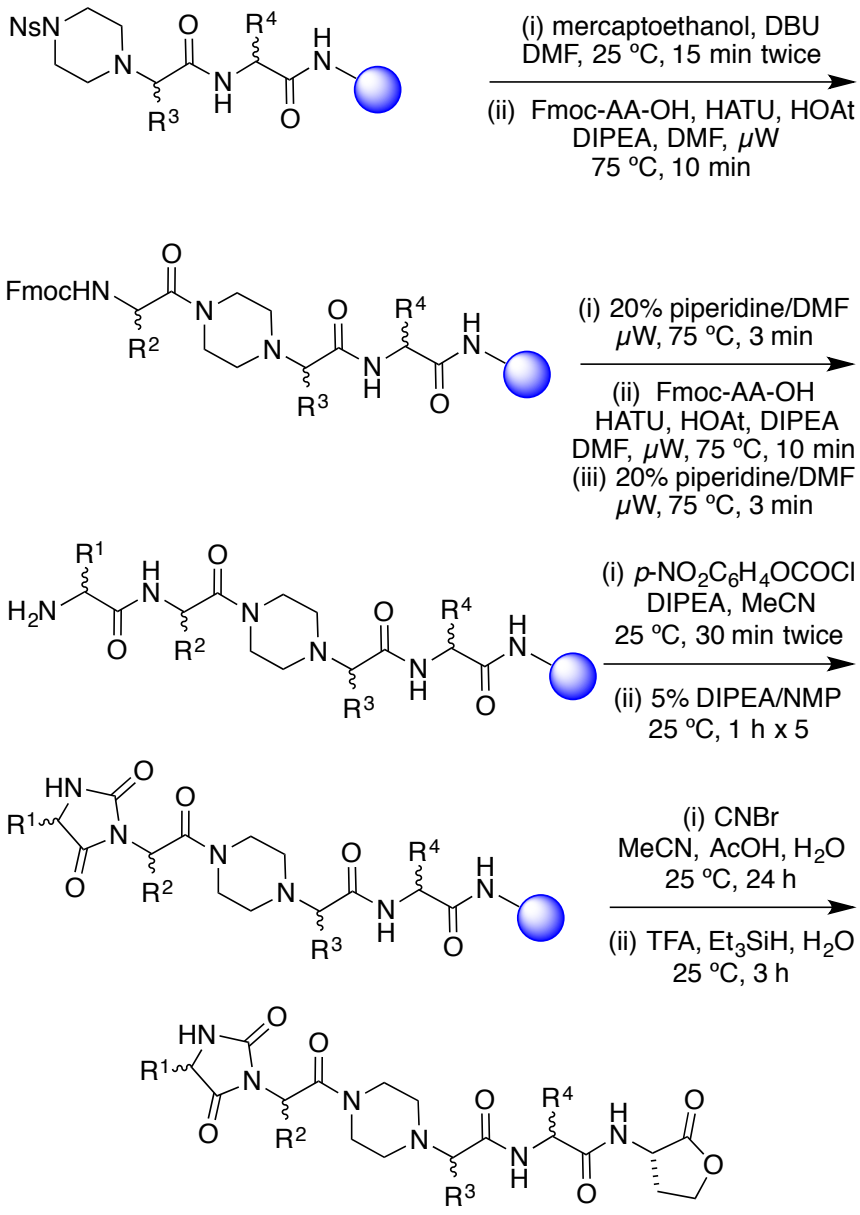
Solid phase syntheses of chemotypes **16** on TentaGel-NH_2 resin were developed (Scheme 4.2). Overall syntheses were similar to the first generation described in chapter III with some modifications. Again, lower case one-letter codes depicted the amino acid side-chains R^1-R^4 as described in previous chapter, and side-chain protecting groups were acid-labile. Methionine linker was attached on TentaGel-NH_2 as

a cleaving group at the end of syntheses. Amino acids R^4 were coupled followed by nosyl-protected piperazine fragments **15** under microwave irradiations. Nosyl-removal as indicated in chapter III,[177] then two more coupling-deprotection cycles using standard *N*-Fmoc-protected amino acids assembled the structures bearing the R^2 and R^1 side-chains. *N*-Fmoc groups were removed, and then converted to hydantoins via a two-step process.[107] Finally, cyanogen bromide was used to cleave the protected chemotypes **16** from the resin, giving the *C*-terminal lactone appendix in these structures. After preparative HPLC purification, compounds were treated with the TFA cocktail to deprotect acid-labile side-chains, and then lyophilized to obtain desired compounds **16**.

Implementation of Scheme 4.2 gave 15 compounds for screening, but one, LLLD-**16ldnq**, was surprisingly vulnerable to degradation in the air, and was not considered further. One of the compounds that was considered, LDLL-**16dl(CN)r**, is a byproduct formed via dehydration of the Asn side-chain in the cyanogen bromide cleavage step.



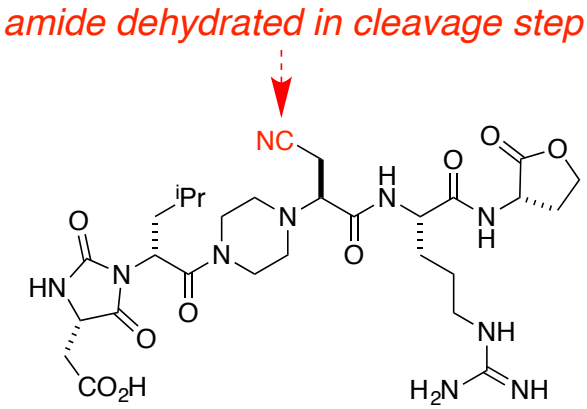
Scheme 4.2 Solid phase syntheses of chemotypes **16** and the structure of LDLL-**16dl(CN)r.**



16

DLLD-**16nclk**; DLLD-**16nclr**; DDDL-**16vclr**
DDDD-**16vclr**; LDLL-**16dlnr**; LDLL-**16dl(CN)r**
LLLD-**16qndw**; LLLD-**16qndr**; DLLL-**16qndr**
LLDL-**16qndr**; LLDD-**16qndr**; DLDL-**16endr**
DLDD-**16endh**; LLLD-**16ldnq**; LLLL-**16aaar**

Scheme 4.2 Continued.



LDLL-**16dl(CN)r**

Scheme 4.2 Continued.

4.4 Binding Assays

All featured second-generation hydantoin-piperazine **16** were tested for their activities to disrupt PCSK9•LDLR interaction. Direct binding experiments were performed by three experiments. Surface plasmon resonance (SPR) was conducted by our collaborator, Dr Xiaowen Liang, from the Center for Infectious and Inflammatory Diseases, Institute of Biosciences and Technology. This method featured measurements of direct binding of inhibitors to the immobilized PCSK9 on a Biacore sensor chip, which could screen several inhibitors in short period and ultimately obtain dissociation constant (K_d). We received PCSK9 protein from Dr. Jay Horton at UT Southwestern Medical Center. Another experiment involved quantification of inhibition using a commercial PCSK9•LDLR *in vitro* binding assay kit (MBL Int. Co). This ELISA assay allowed determining direct inhibition of small molecules to the recombinant PCSK9 and EGF-AB of LDLR, i.e. this assay proved the specific disruption between those two proteins. Photoaffinity labeling was the third experiment to prove that our lead inhibitor bound on the PCSK9. We incubated PCSK9 with the modified lead compound, LDLL-**18dlnr**, irradiated with UV flashlight, and then 'clicked' with azide-fluor 488. This experiment,

including the synthesis of labeling fragment **19**, was conducted by Bosheng Zhou from Dr. Burgess' lab.

4.4.1 Surface Plasmon Resonance (SPR)

Surface plasmon resonance (SPR) was used to screen 14 compounds indicated in Figure 4.2a. The positive control, Pep2-8, was also used in this study, and throughout other experiments. This 13-residue oligopeptide was reported to bind PCSK9 with K_d 0.66 ± 0.11 μM.[133] One compound, LLLL-**16aaar**, is a "partial negative control" having the same core as the EKO-implicated compounds, but only Ala side-chains and an all-L stereochemistry that was not predicted via EKO, i.e. LLLL-**16aaar** controls for random stereochemistry and lacks of functional side-chains. In the event, three compounds and Pep2-8 were selected (on the basis of this initial SPR data and the cellular uptake assays described below) for more thorough SPR analyses to determine dissociation constant (Figure 4.2b-e): LDLL-**16dlnr** (K_d 24.8 ± 9.1 μM), DLDD-**16nclk** (41.2 ± 17.5 μM), LDLL-**16dl(CN)r** (35.8 ± 11.4 μM) and Pep2-8 (3.56 ± 0.16 μM).

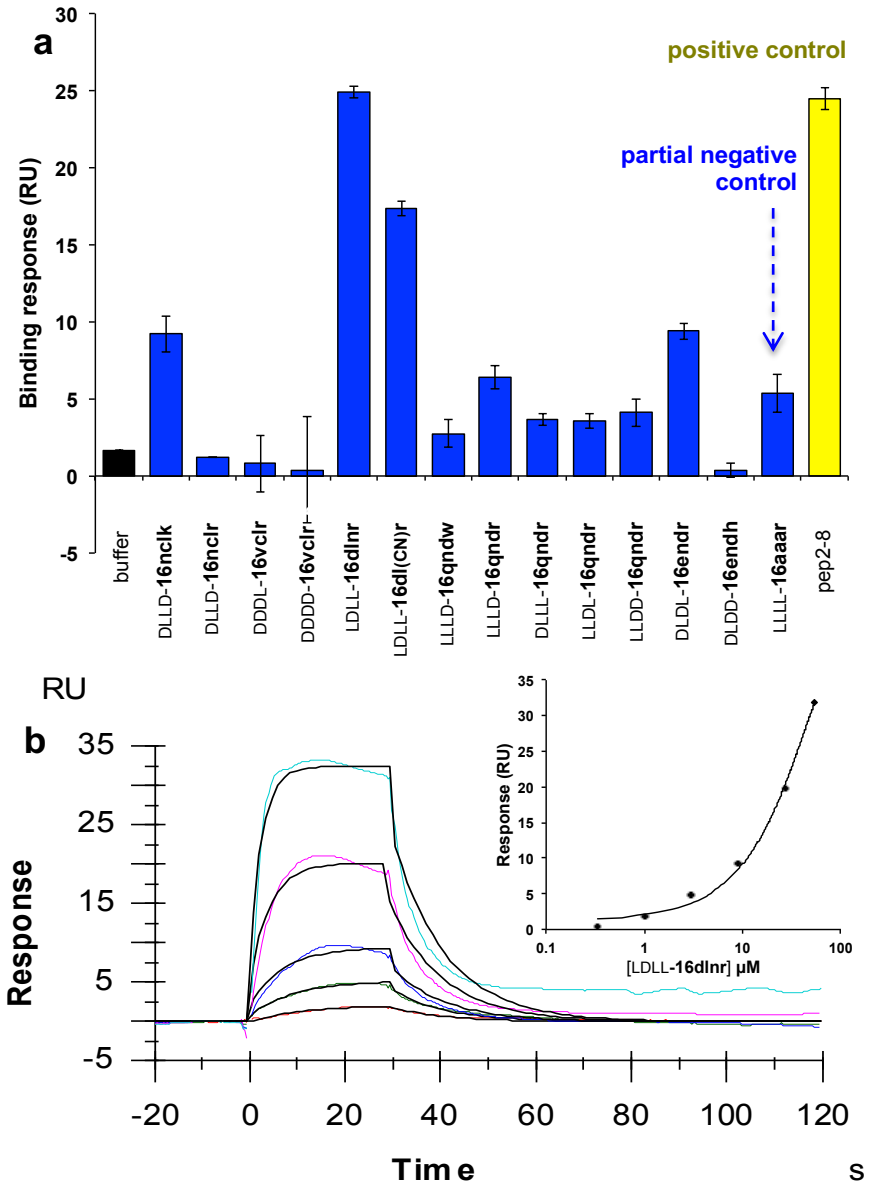


Figure 4.2 (**a**) Initial SPR screening of 14 compounds at 50 µM and Pep2-8 at 5 µM over PCSK9 supported on gold via amine coupling chemistry. (**b-e**) Sensorgram profiles of LDLL-**16dlnr**, DLLD-**16nclk**, LDLL-**16dl(CN)r** (1, 3, 9, 27 and 54 µM) and Pep2-8 (1, 2, 4, 8 and 16 µM) respectively.

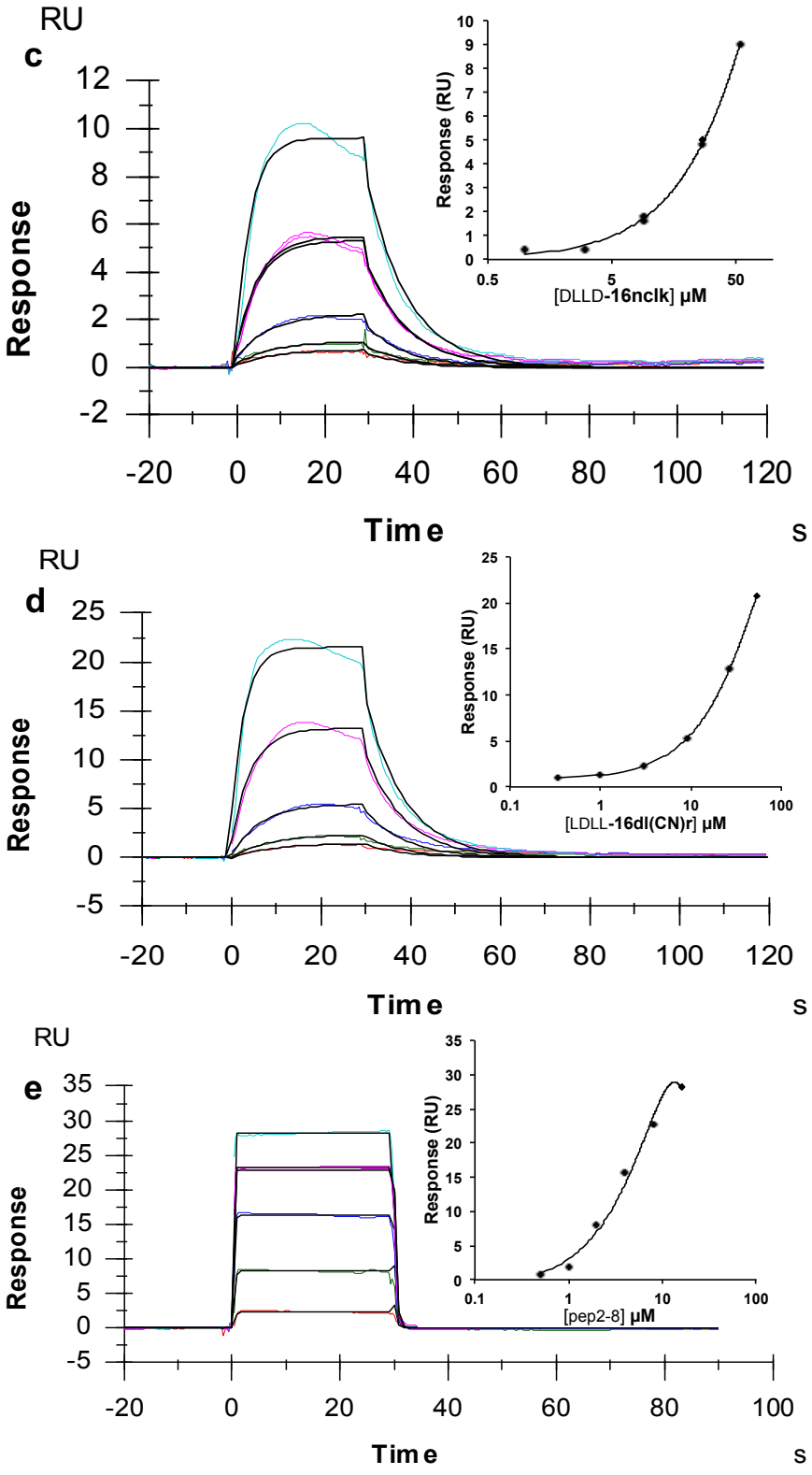


Figure 4.2 Continued.

4.4.2 Enzyme-linked Immunosorbent Assay (ELISA)

Seven compounds in the library (again, selected on the basis of the SPR data and the cellular assays described below) were subjected to an ELISA assay to obtain additional evidence that the small molecules bind PCSK9 (Figure 4.3). In this assay, EGF-AB domain of LDLR was pre-coated in a 96-well plate. His-tagged PCSK9 and inhibitors were incubated together, and then added into the 96-well plate. Inhibitors binding PCSK9 should interfere with PCSK9•LDLR interaction, which in turn reduces the intensities at 450 nm after adding His-tag mAb conjugated with HRP and substrate TMB.

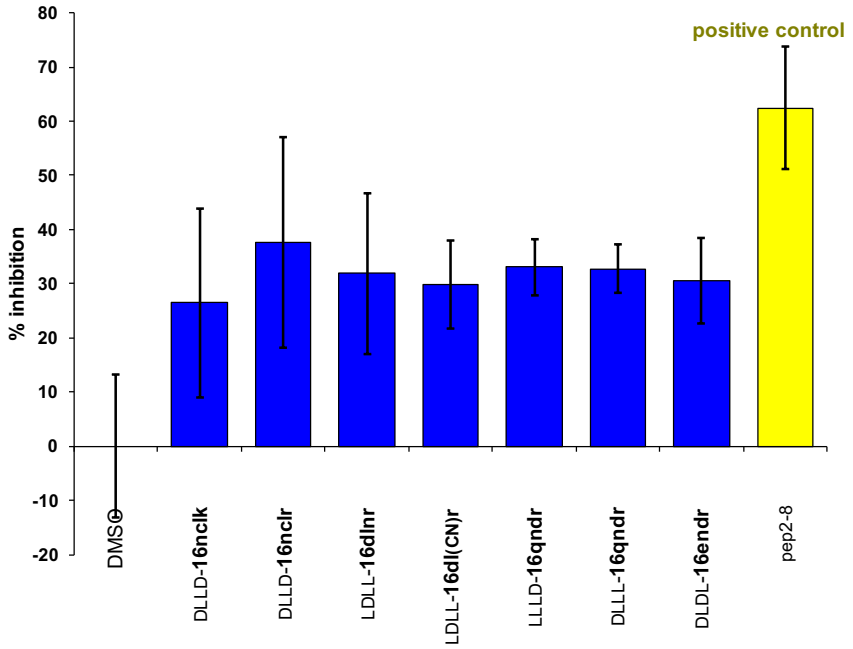


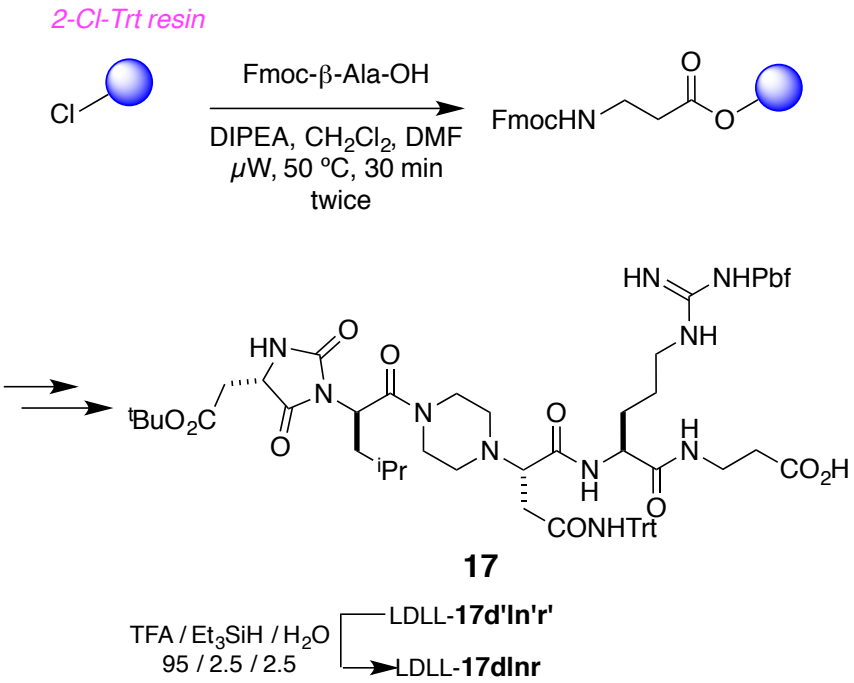
Figure 4.3 PCSK9•LDLR ELISA assay of selected compounds and Pep2-8 (50 μM) with 50 ng/mL PCSK9. Results are the averages ± SD of three independent experiments.

Error limits in this assay are higher than in the SPR experiments. These seven compounds tested showed significant inhibition of PCSK9 to the EGF-AB domain of

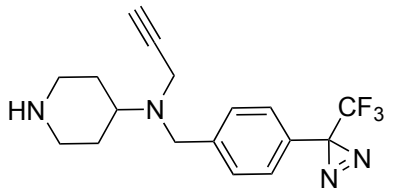
LDLR relative to a negative control that contained both proteins with DMSO (represent complete interaction).

4.4.3 Photoaffinity Labeling

Two derivatives of LDLL-**16dlnr**, compounds **17** and **18**, were prepared to explore if binding of these to PCSK9 could be detected via photoaffinity labeling. Thus, a protected LDLL-**16dlnr** that replaced lactone with β-alanine was prepared on chlorotrityl resin, cleaved with the protecting groups in place, and coupled to a photoaffinity fragment **19** to obtain **18** (Scheme 4.3). Meanwhile, compound **17** without labeling fragment was used as a competitor to **18**.



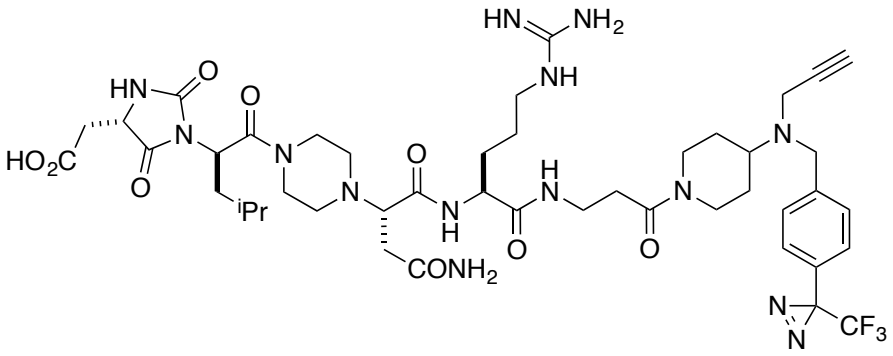
Scheme 4.3 Synthesis of compound **17** and **18**.



19

(i) HBTU, DIPEA, DMF, 0-25 ºC, 12 h

(ii) TFA / Et_3SiH / H_2O (95 / 2.5 / 2.5)



18

LDLL-**18dlnr**

Scheme 4.3 Continued.

Pre-incubation of PCSK9 with **18** and (optionally) with a large excess of the blocking ligand **17**, irradiation of some wells at 365 nm, copper-mediated click reaction with Alexa-488-azide, then SDS-PAGE gave the data shown in Figure 4.4. A fluorescent band corresponding to the molecular mass of labeled PCSK9 (~60 kDa) was observed only in the wells that were irradiated in the absence of the blocking ligand **17** (lane 1, 4 and 7).

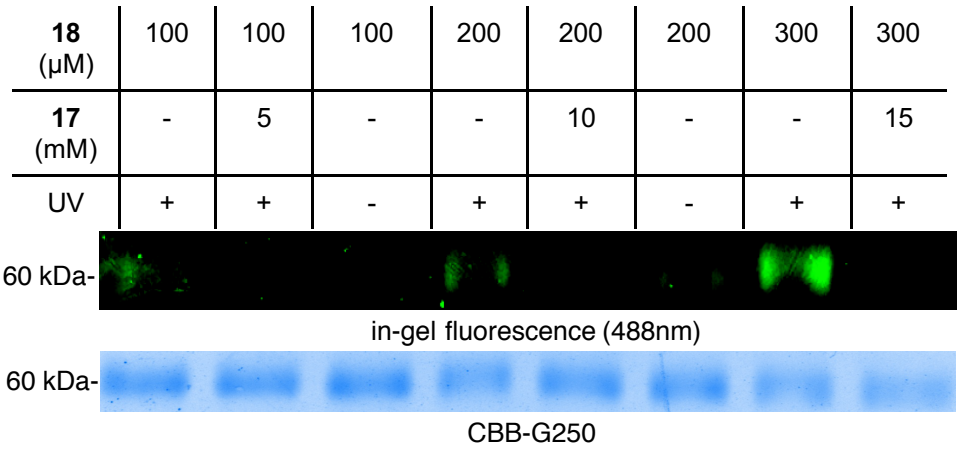


Figure 4.4 Photoaffinity labeling of human PCSK9 protein with compound **18**.

4.5 Biological Assays

Apart from binding assays, we examined our inhibitors in cell culture. Human hepatocellular carcinoma (HepG2) was chosen for this study, which was reported to have abundant LDLRs on cell surface, and was a standard cell line used in the study of PCSK9.[121,130,133,162,178] Two featured experiments, LDL-uptake and cell-surface LDLR assays, were performed to determine the biological actives of small molecules toward live cells.

Before evaluating our compounds, seven selected compounds shown in Figure 4.3 were tested for cytotoxicity using liver cells (HepG2). All seven of these compounds showed no toxicity up to 50 μM. The compound with the lowest K_d in SPR studies, LDLL-**16dlnr**, was also checked at 100 μM and showed no cytotoxicity even at this concentration (Figure D.1). Another preliminary experiment to the key cellular assays, the water solubilities of the featured compounds were measured. The solubility concentration

gradients for these materials were linear to beyond 100 μM, i.e. they were soluble at the maximum concentration used in the uptake assays below (Figure D.2).

4.5.1 LDL-uptake Assay

An established assay for PCSK9•LDLR inhibition features uptake of fluorescently labeled LDL nanoparticles (BODIPY-LDL, Invitrogen) by hepatocytes; the cells become fluorescent as the particles are absorbed.[133] Uptake of the BODIPY-containing particles is maximized in the absence of PCSK9, and the cells become fluorescent; conversely, adding PCSK9 diminishes that signal. Addition of PCSK9 *and* a compound that interferes with the PCSK9•LDLR interaction would be expected to give cells that are more fluorescent than those to which only PCSK9 was added but less so than cells to which none of that protein was present.

Figure 4.5a shows maximal uptake (calibrated to 100 %; black bar) in the absence of PCSK9, while all the other data points correspond to 15 μg/mL of that protein; the "negative" corresponds to only PCSK9 added (calibrated to 0 %). Pep2-8 at 30 μM restored the LDL uptake (yellow bar) to within 80 % of its maximal value (black). Several of the featured chemotypes **16** showed promise insofar as they, like Pep2-8, also restored fluorescence; the ones marked with a red dagger were selected for further assays on the basis of this data and the SPR studies above. Recall that LLLL-**16aaar** is a "partial control" as described above (same chemotype, just methyl side-chains); it did *not* induce significant BODIPY-LDL uptake.

Figure 4.5b shows data derived from repetition of these experiments under identical conditions except that three different doses of the test compounds were used. Overall, all the compounds show a dose response, except LLLD-**16qndr**, DLLL-**16qndr**, and, as expected, the partial control LLLL-**16aaar**. Figure 4.5c shows a more extensive dose-response curve for one of these compounds, LDLL-**16dlnr** (again, selected on the basis of the overall data); this data shows an encouraging correspondence.

It is curious that LDL uptake in the HepG2 cells was significantly enhanced when 100 μM of compound was used (Figure 4.5c). This is consistent with the SPR binding data in which LDLL-**16dlnr** showed a longer resident time (15-20 fold slower off-rate) compared with Pep2-8 (Table D.1). More particularly, the longer half-life of the LDLL-**16dlnr**•PCSK9 complex became more obvious when higher concentration (100 μM) of compound was injected onto the PCSK9-functionalized surface. One explanation for these observations is that there could be a synergistic target site for LDLL-**16dlnr** that only becomes significant at higher compound concentrations.

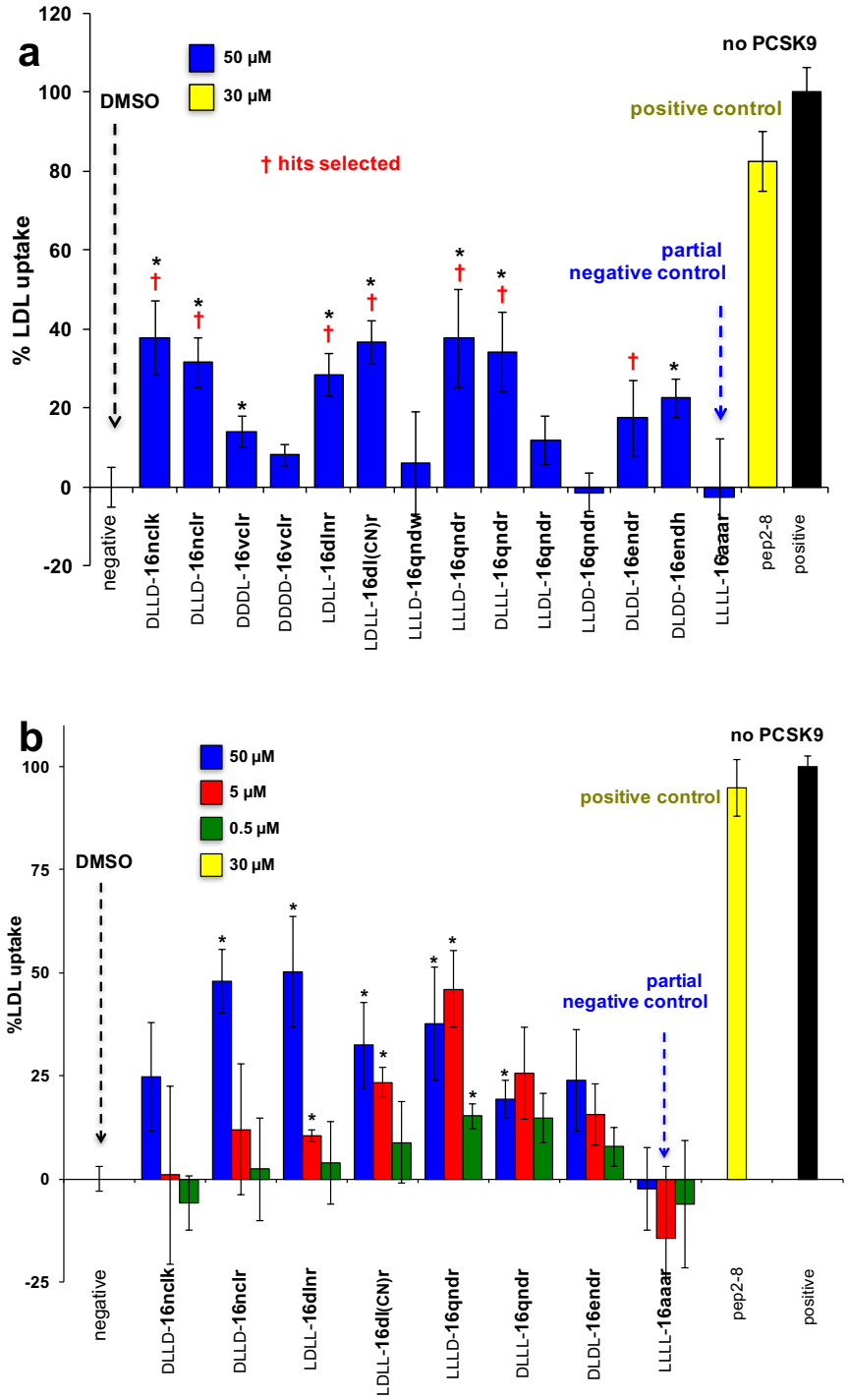


Figure 4.5 Uptake of BODIPY-LDL by hepatocytes. (**a**) Initial screen at 50 µM concentrations. (**b**) Data of select compounds at three different doses. (**c**) A more extensive dose-response curve of one select lead: LDLL-**16dlnr**. All results are represented as means ± SD of three independent experiments. Significant differences

between compounds and negative control are determined using student' t-test (* $p \leq 0.05$, ** $p \leq 0.01$).

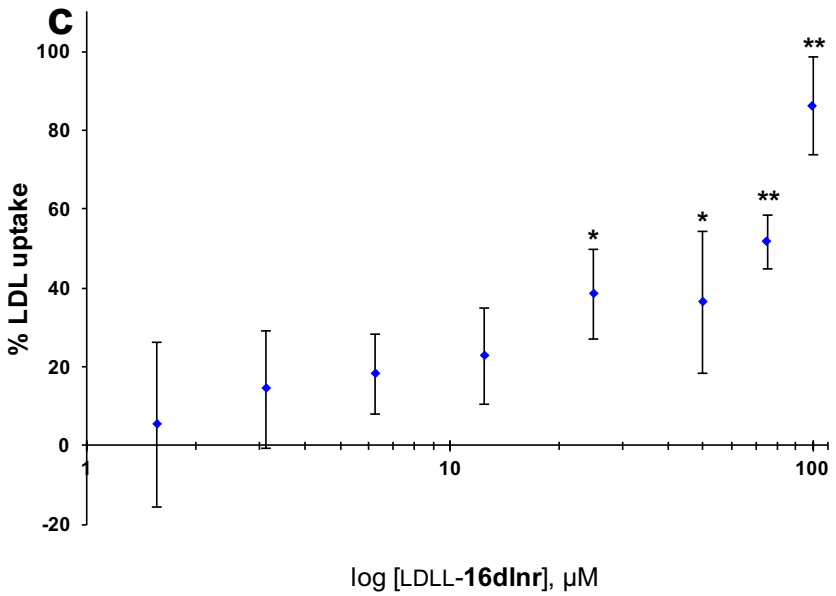


Figure 4.5 Continued.

4.5.2 Cell-surface LDLR Assay

Two assays were attempted to probe if increased uptake of LDL particles correlates with increased expression of LDLR. Our initial attempts to do this featured monitoring LDLR levels on treated and untreated hepatocytes using flow cytometry. HepG2 treated with PCSK9 would display the lowest level of LDLR, whereas cells treated with inhibitors or without PCSK9 would have higher LDLR levels. The data obtained (Figure 4.6) indicated an increase in LDLR levels upon treatment with the seven hit compounds (as shown in Figure 4.5b), but the errors in the measurements were such that the increases had borderline statistical significance.

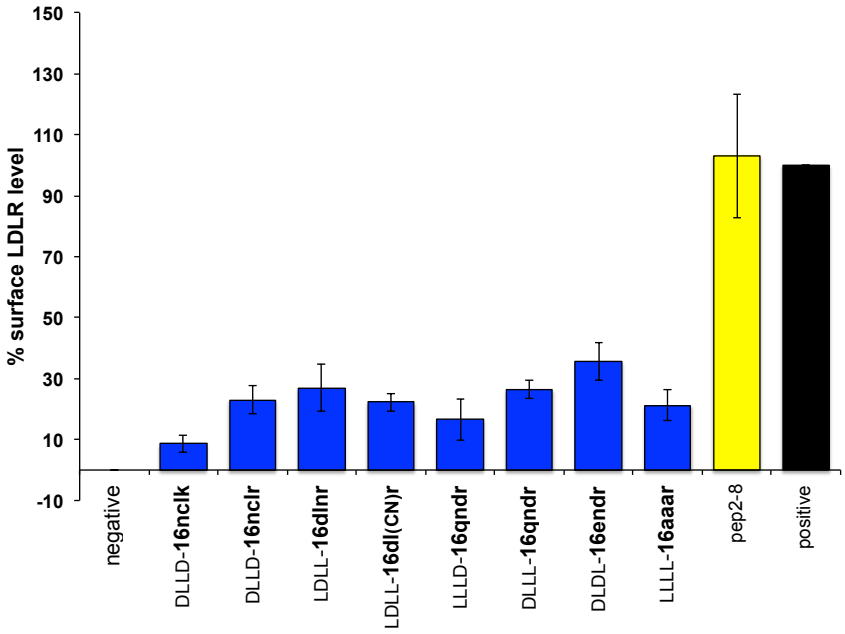


Figure 4.6 Percent surface LDLR level on the surface of HepG2 cells.

Consequently, we resorted to a semi-quantitative approach in which the LDLRs on live hepatocytes (treated and untreated) were visualized using an LDLR-selective mAb in combination with an Alexa Fluor®-labeled secondary mAb. Figure 4.7a shows expression of LDLR in the cells was suppressed when they were treated with PCSK9 alone. When cells were treated with the test compounds, then they stained more brightly (Figure 4.7b-e) though not as brilliant as the Pep2-8 or positive culture that did not incubate with PCSK9 (Figure 4.7f-g). Figure 4.7h plotted the average corrected total cell fluorescence (CTCF) of Figure 4.7a-g by ImageJ software. All three selected compounds (DLDD-**16nclk,** LDLL-**16dlnr,** DLLL-**16qndr**) had average fluorescent intensity higher than the partial negative control LLLL-**16aaar**, which supported the observation from LDL-uptake assay. In the process of fluorescent quantification, we *manually* selected all cells in each image. These measurements were estimated, so that

the error bars were not reported. Nevertheless, the conclusion was still valid; the featured compounds were able to disrupt PCSK9•LDLR interaction.

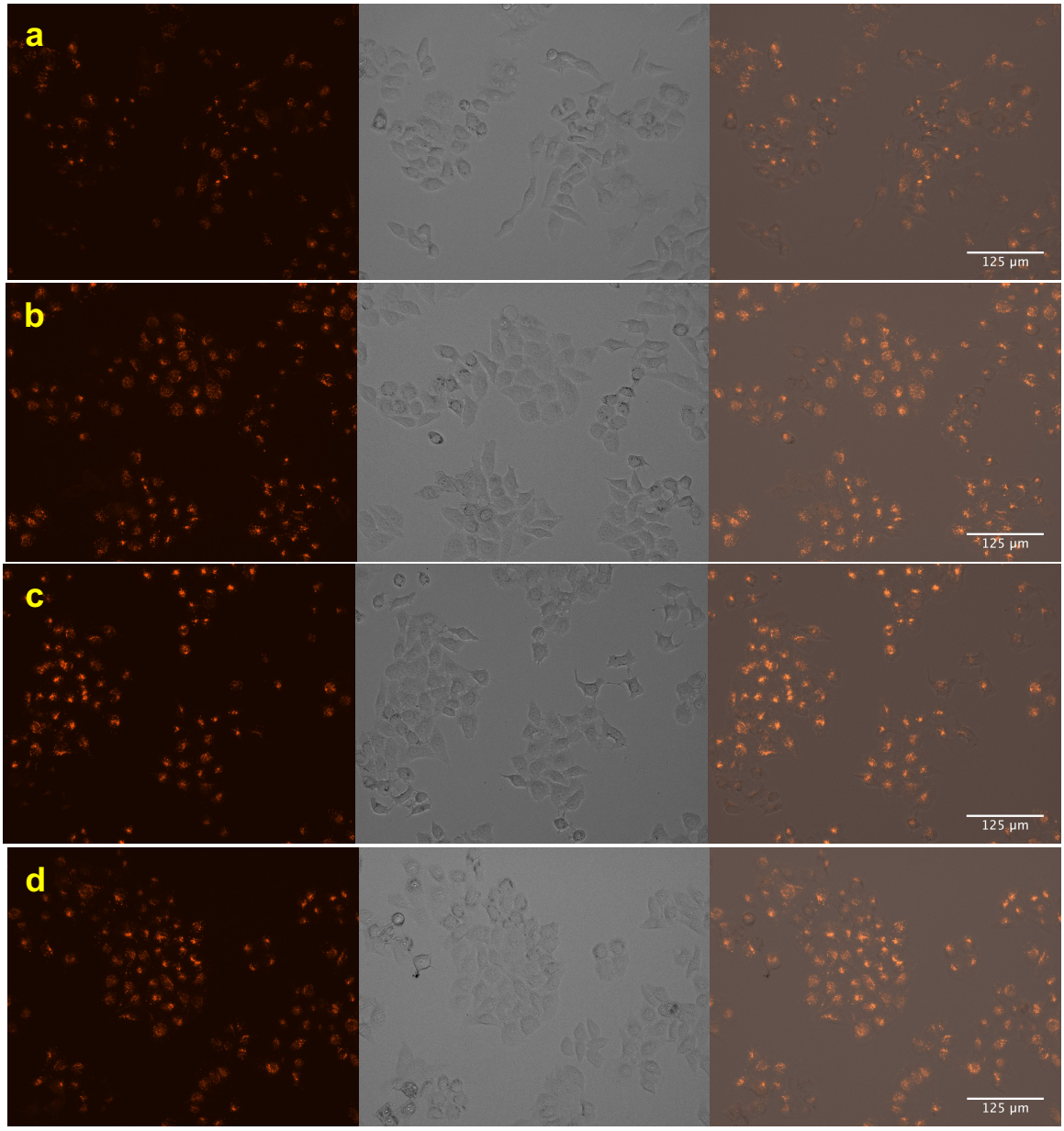
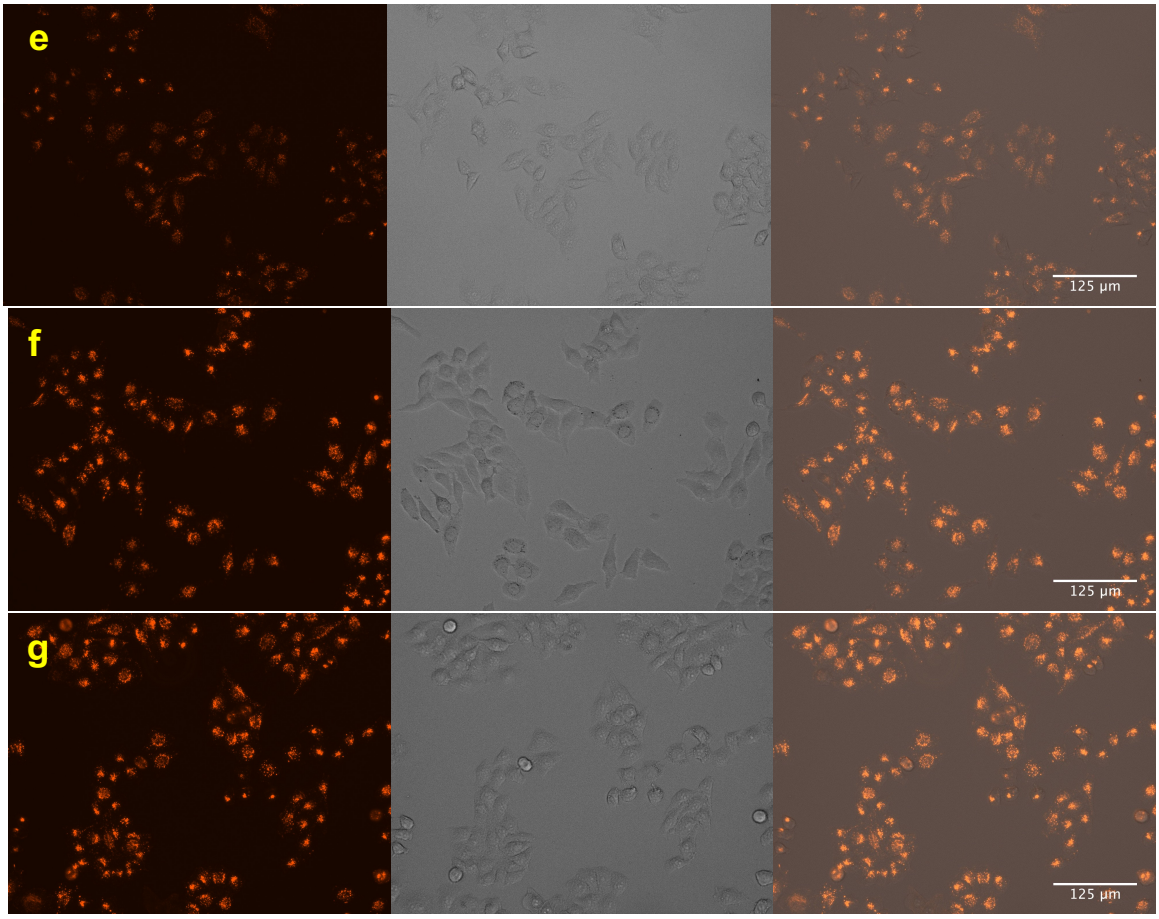


Figure 4.7 Fluorescent imaging of HepG2 cells treated with PCSK9 and (**a**) DMSO; (**b**) DLDD-**16nclr**; (**c**) LDLL-**16dlnr**; (**d**) DLLL-**16qndr**; (**e**) LLLL-**16aaar**; (**f**) Pep2-8; and (**g**) DMSO without PCSK9. (**h**) Fluorescent intensity was quantified by ImageJ software.



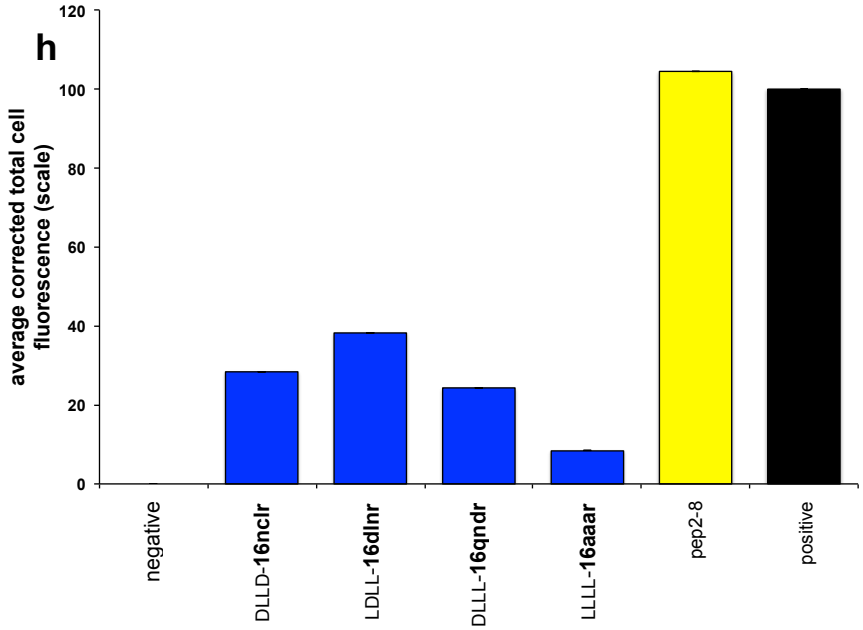
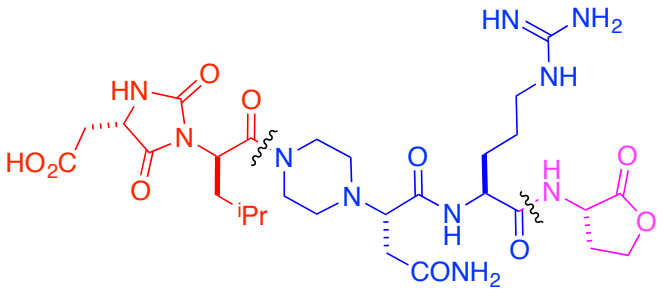


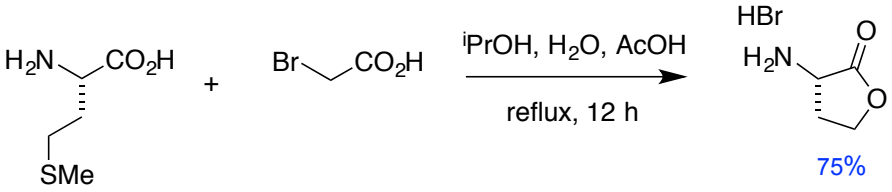
Figure 4.7 Continued.

4.6 Syntheses of LDLL-16dlnr Mimic via Solution-phase Syntheses

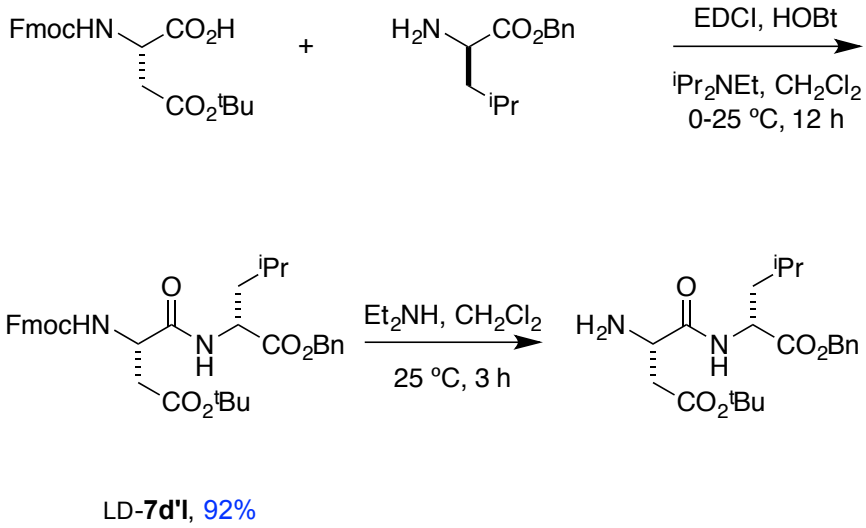
The results from binding and biological assays indicated the LDLL-**16dlnr** was a promising lead, so we decided to scale up the reaction via solution phase synthesis to prove accessibility. LDLL-**16dlnr** can break down into 3 fragments, hydantoin, dipeptides and homoserine lactone. The outline syntheses are shown in Scheme 4.4. L-homoserine lactone was prepared from L-methionine and bromoacetic acid according to the literature.[179] This lactone was coupled with *N*-Fmoc-arginine, and subsequently nosyl-protected piperazine-asparagine using HBTU activator. Deprotection conditions of nosyl group were screened with various reagents, and the best was the same used in solid-phase synthesis (mercaptoethanol/DBU). Hydantoin ring was synthesized from Asp-Leu dipeptide. After Fmoc deprotection, dipeptide was treated with *p*-nitrophenyl chloroformate in organic solvent, followed by addition of water to cyclization. We observed the epimerization if organic base (e.g. Et_3N) was used instead of water. Finally, benzyl was hydrogenolyzed, and the hydantoin was coupled with piperazine fragment to obtain protected form of LDLL-**16dlnr**. Overall, we successfully prepared **16** in gram scale via solution phase synthesis.



a

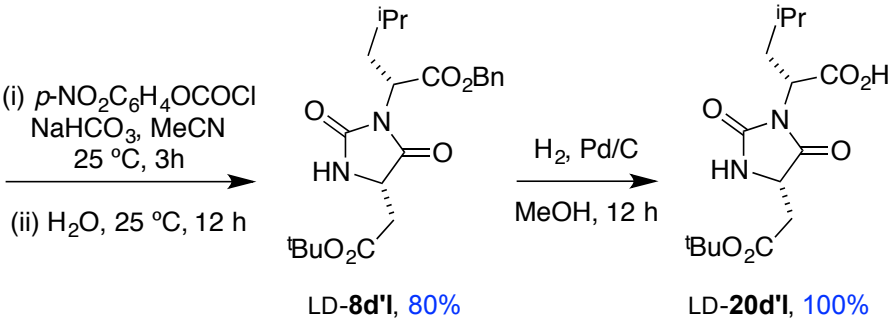


b



LD-**7d'l**, 92%

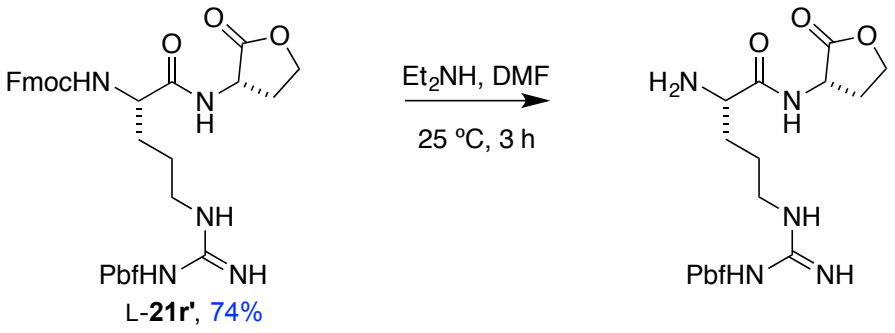
Scheme 4.4 Syntheses of LDLL**-16dlnr** in solution phase.



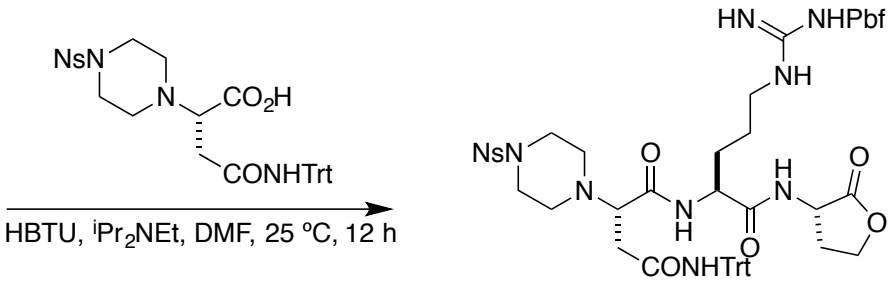
LD-**8d'l**, 80%

LD-**20d'l**, 100%

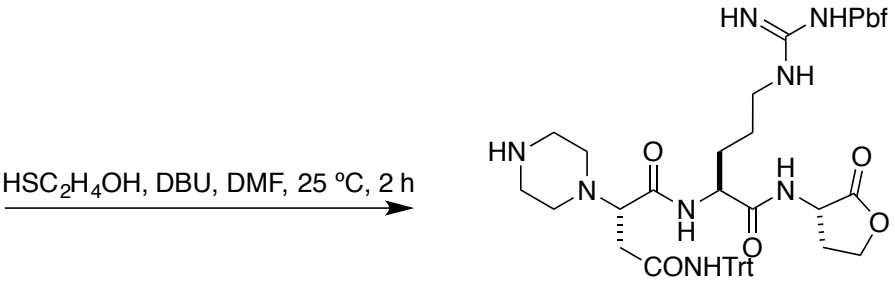
c



L-**21r'**, 74%

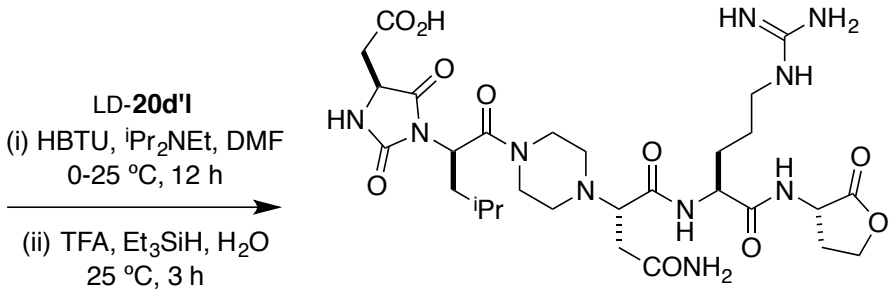


LL-**22n'r'**, 65%



LL-**23n'r'**, 96%

Scheme 4.4 Continued.



(i) LDLL-**16d'ln'r'**, 77%
(ii) LDLL-**16dlnr**, 100%

Scheme 4.4 Continued.

4.7 Conclusion

Compounds **12** implicated by EKO approach had weak binding affinity to be detectable in the TR-FRET experiment, as described in Chapter III. To overcome this issue, the structural modifications with docking programs are desired. We chose Glide in this work because (i) it allowed users to build their own libraries; (ii) docking parameters were adjustable to imitate ligand poses from EKO; and (iii) docking accuracy was reliable comparing with other similar docking programs.[176] By using eight compounds **12** as parental templates, we substituted amino-acid lactones at the *C*-termini to conceive fifteen second-generation hydantoin-piperazine mimics **16**. Experiments from both binding and cellular assays revealed seven out of fifteen mimics **16** showed significant, measureable activities against PCSK9•LDLR interaction. Further evaluation of those seven compounds indicated that LDLL-**16dlnr** was the lead among the second-generation mimics with K_d = 24.8 ± 9.1 μM on PCSK9 based on SPR experiment, and approximately 50% increase of LDL uptake at 50 μM. These results were in agreement with the photoaffinity data wherein the modified structure, LDLL-**18dlnr**, bound on the PCSK9 protein, and the ability of LDLL-**16dlnr** to restore LDLR display on HepG2 cells

observed by fluorescent imaging. We also synthesized LDLL-**16dlnr** in solution phase to prove the syntheses of this analogue can be scaled up if desired. Apart from this lead, DLDD-**16nclk**, LDLL-**16dl(CN)r** and LLLD-**16qndr** were also effective to disrupt PCSK9•LDLR interaction.

There are opportunities for structural modifications to further improve binding affinities and increase LDL uptake characteristics. Homoserine lactone is an appendix from cyanogen bromide cleavage, in which this lactone was not in our original strategy to improve binding affinities of mimics **12**. Indirect evidence from LDLL-**18dlnr** implied that this lactone might not significantly contribute the binding energy to PCSK9. Moreover, it is reasonable to synthesize molecules bearing R^1-R^3 side-chains same as LDLR side-chain sequences to mimic natural interaction between PCSK9•LDLR, thus those molecules are likely to dock on PCSK9 selectively. However, there is a possibility to gain superior binding interactions with different side-chains apart from natural interaction, i.e. those mimics may fit well in the pockets of PCSK9 protein. In this regard, we can explore the side-chain combinations that provide better docking score via CombiGlide. Ultimately, the results from this project are a stepping stone to make better potent small molecules to disrupt PCSK9•LDLR interaction.

CHAPTER V

EVALUATION OF SECONDARY STRUCTURE MIMICRY TO DISRUPT PROTEIN-PROTEIN INTERFACES

5.1 Introduction

Minimalist mimics are compounds that do not have peptidic backbones, but nevertheless preferentially project amino acid side-chains in orientations that resemble secondary structures. Decades of research in this area has been driven by the assumption that if a secondary structure is found at a protein-protein interface then a corresponding minimalist mimic is a candidate to displace the protein it mimics from the interface,[10,11,26] and there are several instances in which minimalist mimics have been proven to disrupt PPIs (examples[41-46] and reviews[40,47-51]). However, *the existence of a relationship between secondary structure- and interface-mimicry is an untested hypothesis; it cannot be confirmed on the basis of the experimental data available since there are not enough data points.* Consequently, even though secondary structure mimicry is widely seen as a fast-track to molecules that disrupt specific PPIs, the value of secondary structure mimicry is assumed rather than proven. We saw an opportunity to test the hypothesis that secondary structure mimicry is important using a combination of four computational approaches: EKOS, EKO, DSSP, and STRIDE.

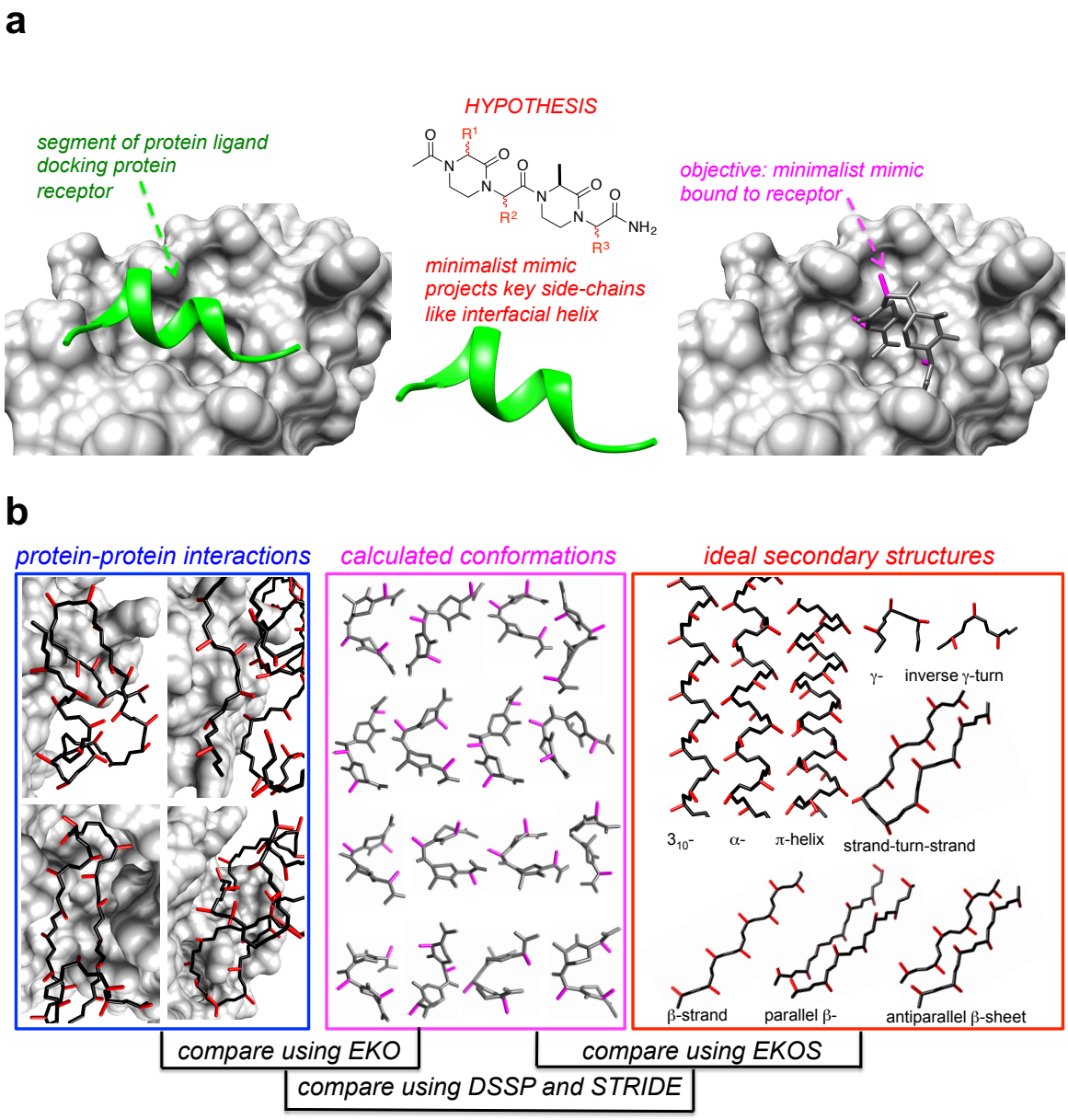


Figure 5.1 (**a**) The secondary structure mimic hypothesis involves (top row): (left) identification of a PPI with an interfacial secondary structure motif; (center) identification of a mimic of this secondary structure; and, (right) displacement of the protein having the secondary structure from the PPI. (**b**) Work flow in this study: comparing preferred conformations with protein-protein interfaces (EKO), with ideal secondary structures (EKOS), then the matches in both to detect secondary structures at interfaces (DSSP and STRIDE).

As this research evolved, the validity of matching protein interface residues to secondary structures emerged as an important issue. EKO is used to determine PPI interface segments where preferred conformations of the featured chemotypes overlay based on all three $C\alpha$ - $C\beta$ vectors. An open-source computational program called the *Dictionary of Secondary Structures of Proteins* (DSSP) was then used to evaluate the presence of secondary structures at protein-protein interfaces. DSSP identifies secondary structure motifs based on hydrogen-bonding patterns.[180,181] Throughout this work, DSSP was set so that all three $C\alpha$ - $C\beta$ vectors must match on a particular secondary structure otherwise the overlay is called "segment". To verify DSSP data, another program, called "STRIDE", was also used for the same purpose. Like DSSP, STRIDE evaluates protein residues in terms *H*-bonding patterns, but it also uses dihedral-angle parameters.[182,183] Both programs report secondary structure motifs using same one-letter codes (H = α-helix; G = 3_{10}-helix; I = π-helix; E = sheet; B = strand; T = turn; S = bend; _ = uncategorized segment). STRIDE does not differentiate between turn and bend motifs like DSSP, so we combined those two results as turns/bends category. We also combined sheet and strand motifs into the same group since both programs did not classify the type of sheet motifs (β-strand, parallel- or antiparallel-β sheet).

Figure 5.2 illustrates some typical data. All the data in Figure 5.2 relates to helices, though just for illustration; similar considerations apply to other secondary structures. In Figure 5.2a a mimic overlays on a near-ideal α-helical fragment in a protein (at an interface); both DSSP and STRIDE call the region where the conformer overlaid as helical (H). However, in Figure 5.2b a conformer is overlaid on an extended region between two helical segments; in fact DSSP calls the residues overlaid as turn, turn, and helical (TTH) while STRIDE concludes this is helical (HHH). In this work, the DSSP

result is classified as “segment” (not overlaid on and particular secondary structure), while STRIDE “bins” this region as helical. In our view, it is debatable if that extended region should be denoted as helical. Similarly in Figure 5.2c, it is unclear that that extended, twisted region is helical: DSSP would bin that as a segment, while STRIDE calls it as a helix. In Figure 5.2d, the mimic has overlaid on an extremely distorted region between two helical fragments; both DSSP and STRIDE bin this overlay as a segment. Due to the distortion illustrated in these examples, we decided to set that all three $C\alpha$ - $C\beta$ vectors must match on a particular secondary structure otherwise they were called “segment”.

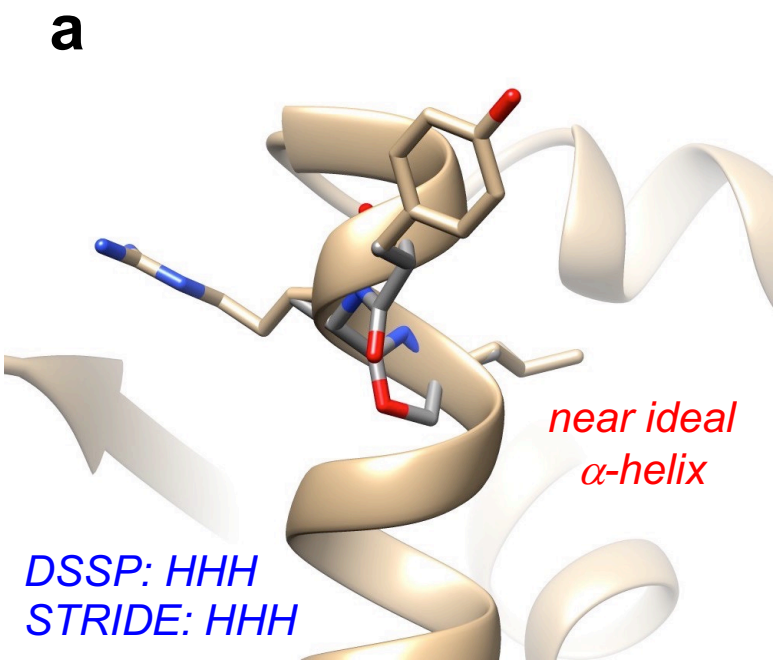


Figure 5.2 Illustrative examples of DSSP and STRIDE secondary structure assignment at protein interfaces where mimics overlay.

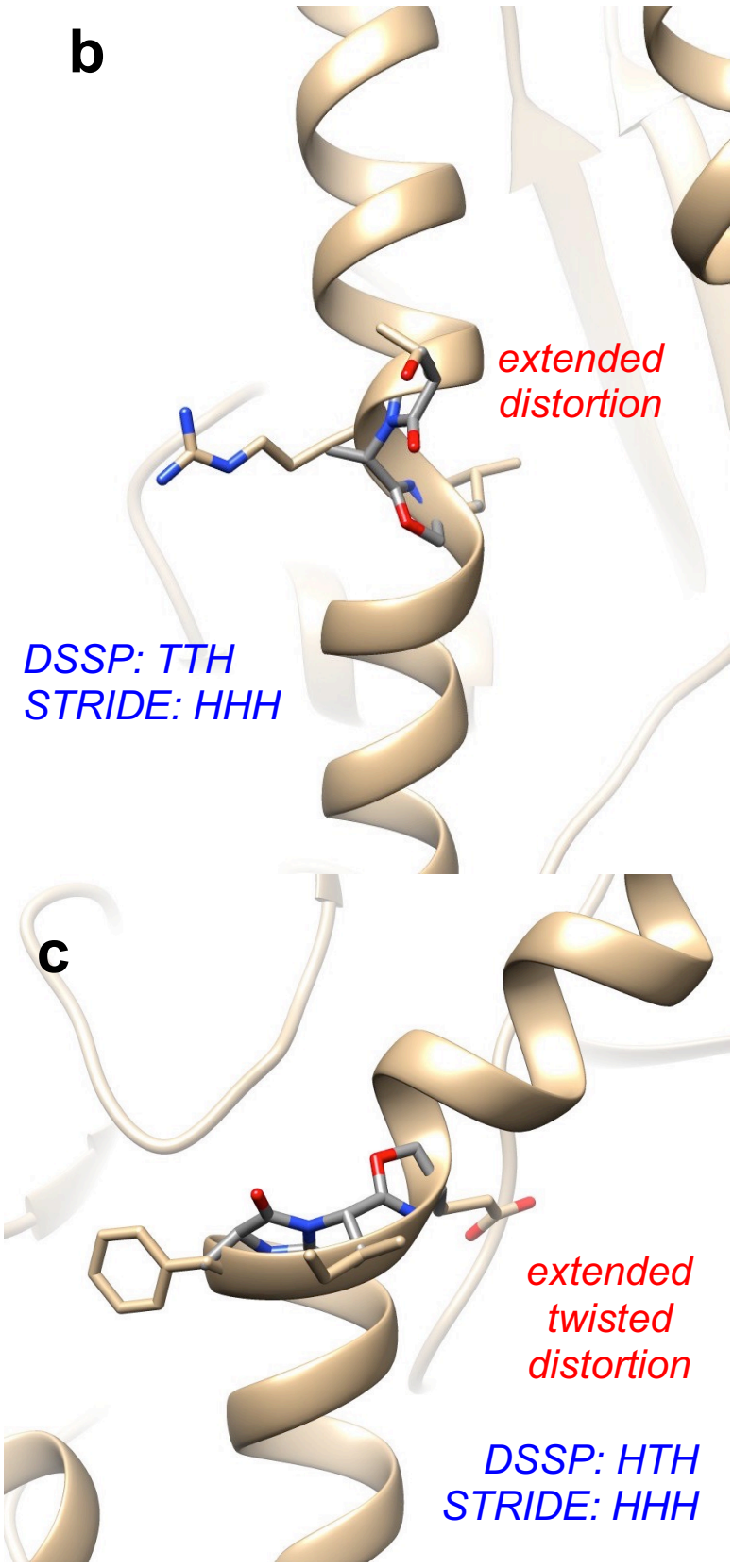


Figure 5.2 Continued.

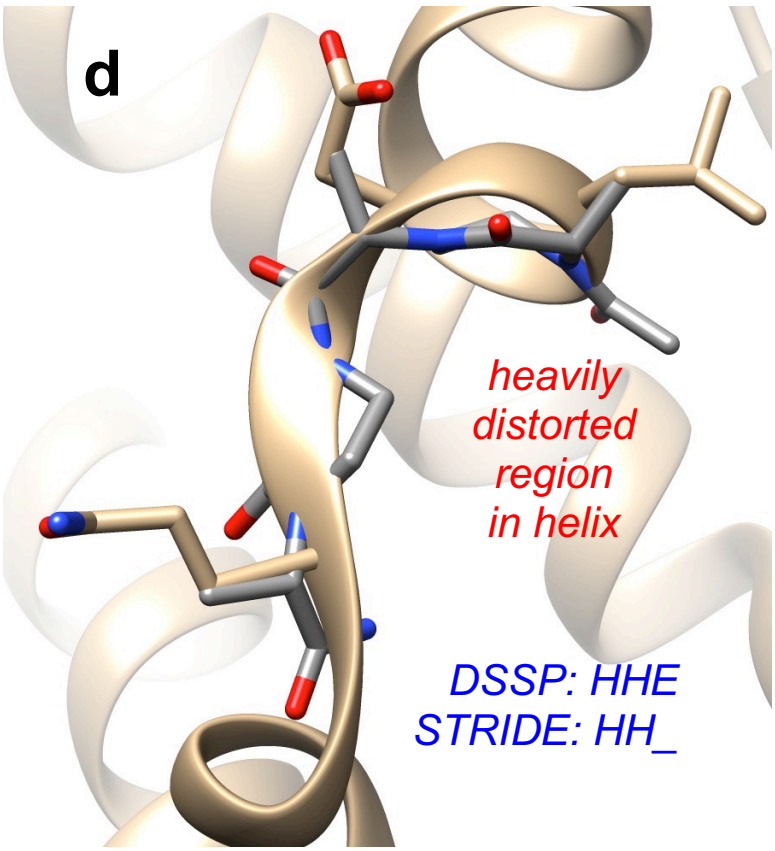


Figure 5.2 Continued.

Only four illustrative examples are given in Figure 5.2, and these do not cover every case wherein distortions may present at PPI interfaces. However, comparison of DSSP and STRIDE data reveals that STRIDE tends to allow more deviation from ideal than DSSP. However, the difference in DSSP and STRIDE outputs is not significant enough to affect the overall conclusions.

In this work we designed chemotypes that may present side-chains in orientations that resemble peptides and proteins without intending to target any particular secondary structure, then asked if their preferred conformers mimic secondary structures and interface regions. Thus the research described here features eight minimalist mimics that have not been prepared or reported previously, and Arora's oxopiperazine

chemotype **A**[38,184] was also included as a reference. Our aims were to: (i) determine how these minimalist mimics compare with huge numbers of interfaces in approximately 240,000 crystallographically characterized PPIs; (ii) compare this data with the bias of the same mimics to adopt conformations that mimic secondary structures; and, (iii) evaluate the presence of secondary structures at the interfaces where the mimics do overlay well (using EKO, EKOS, and DSSP/STRIDE, respectively; Figure 5.1). If secondary structure mimicry is important in the design of interface mimics then chemotypes that overlay well on ideal secondary structures should match frequently at PPI interfaces. A number of conformers and clusters of each chemotype used in these studies are summarized in Table E.1.

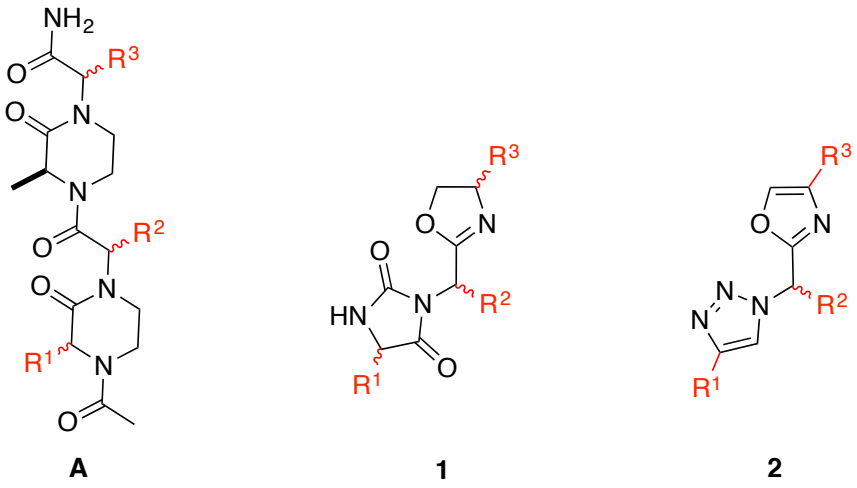


Figure 5.3 Structures of chemotypes **A**, **1**, **2**, **24**-**29**

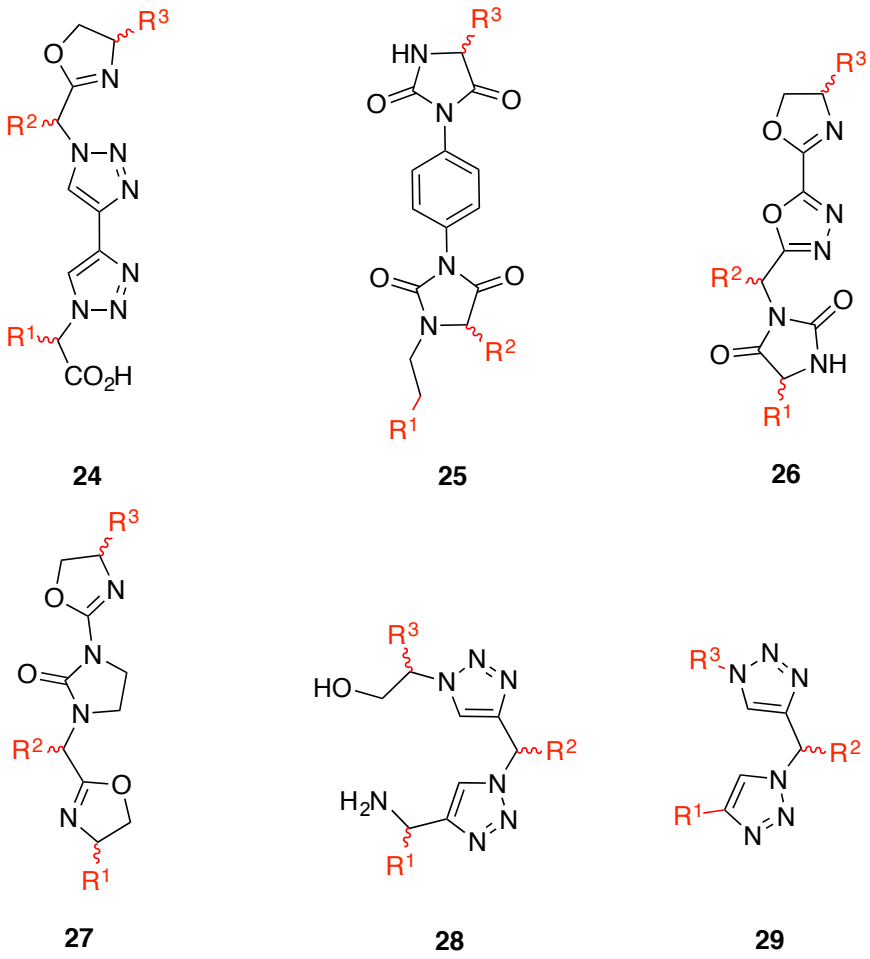


Figure 5.3 Continued.

5.2 Analyses of Peptidomimetics A, 1, 2, 24-29

*5.2.1. Arora's Oxopiperazine **A**: A Benchmark*

EKOS simulations indicate the trimethyl-substituted chemotype LLL-**Aaaa** tends to overlay select common secondary structures with RMSDs smaller than the best-known minimalist mimics as of 2014.[54] Consequently, chemotype **A** is a useful benchmark for good interface mimic design. Figure 5.4a shows how each of the eight possible

stereoisomers (grouped on the x-axis) overlay on the ideal secondary structures, and Figure 5.4b arranges the best matching conformers in descending RMSD of the overlays irrespective of stereochemistry. The best overlay identified was for LDD-**A** on a parallel β-sheet (RMSD 0.21 Å). The second-best overlay had a significantly different RMSD (0.28 Å) indicating the optimal one was somewhat unique. Overall, some conformers of **A** tends to overlay better on extended structures than the helical ones. For any helical structure, the best overlay was for LLL-**A** on the *i, i* + 1, *i* + 3 side-chains of a π-helix (0.36 Å RMSD).

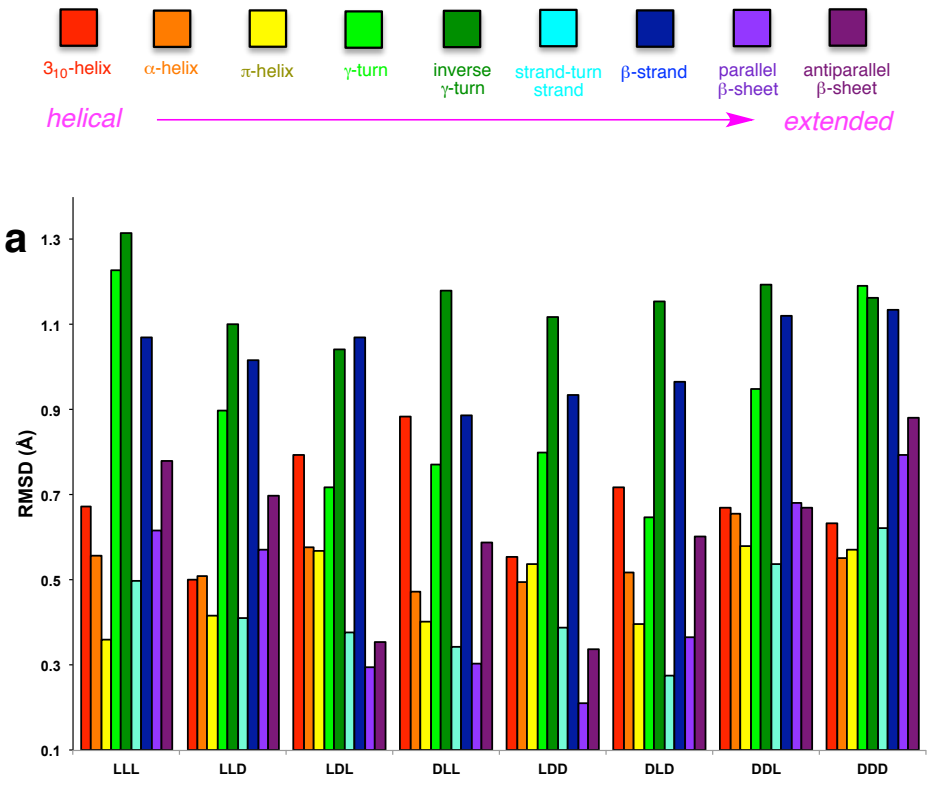
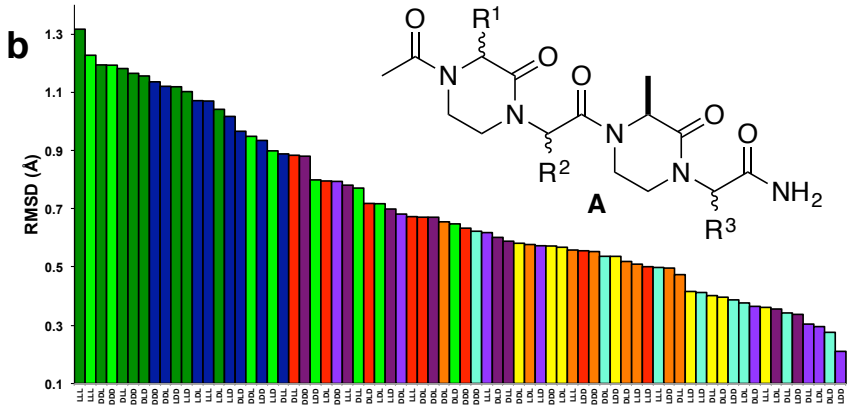


Figure 5.4 RMSD (Å) of the overlays of mimics **A** on each of the ideal secondary structures, organized by stereochemistry (**a**) or by decreasing RMSD (**b**). Statistical distribution of secondary structures at PPI interfaces derived by DSSP and STRIDE calculations; (**c**) the best 312 overlays of LLL-**A** (all RMSDs < 0.25 Å); and, (**d**) 320

overlays of LDD-**A** (RMSD < 0.25 Å). Note that calculations do not differentiate strand-turn-strand, parallel- and antiparallel-sheets.



c

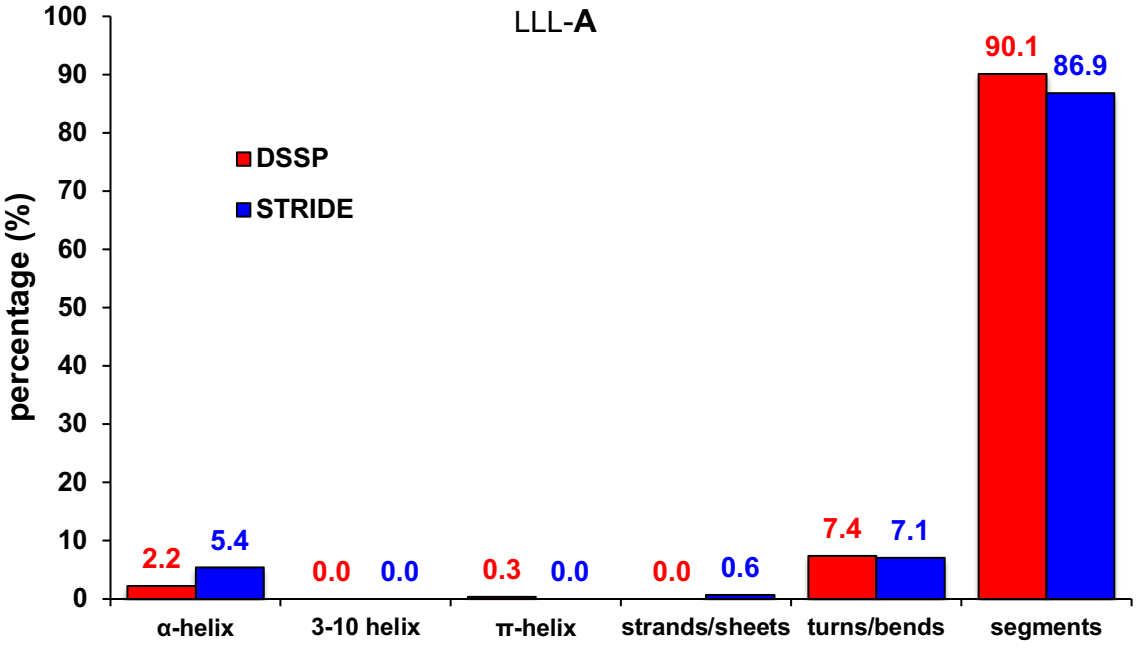


Figure 5.4 Continued.

d

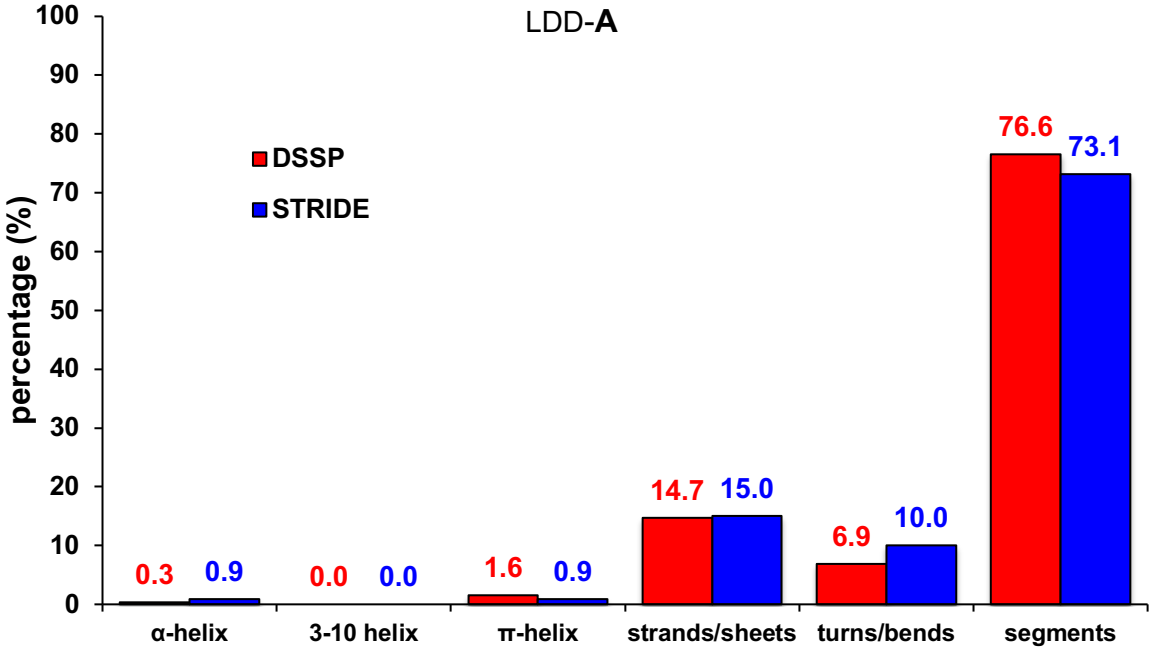


Figure 5.4 Continued.

LLL- and LDD-**A** were selected for a study featuring mining of most of the PPI interfaces in the PDB via EKO. Taking a threshold of 0.25 Å gave 312 and 320 PPI interfaces, respectively. Each match was then analyzed using DSSP and STRIDE. Both calculations were set so that all three *C*α - *C*β vectors must match on a particular secondary structure otherwise the overlay is called "segment". Using these criteria, few of the matches observed corresponded to the mimic overlaid on an ideal secondary structure (Figure 5.4c and d). EKOS analyses on ideal secondary structures predicted the isomer that best matches the secondary structures overall was LDD-**A**, and this favors sheets and turns over helices at interfaces, *but most (>73 %) of the matches were on segments* (Figure 5.4d). Consistent with this, the EKO analysis on PPI interfaces gave optimal matches for sheet regions. However, some stereoisomers like LLL-**A** also

showed biases towards helical conformation (based on EKOS; Figure 5.4a,c), though turns and bends were slightly dominant.

*5.2.2. Hydantoin-oxazoline Chemotype **1**: A Universal Peptidomimetic*

Interface mimic **1** has the same "chain periodicity" as three continuous amino acids (*i*, *i*+1, *i*+2): each side-chain is separated by two atoms in the main-chain (Figure 5.5). In its most stable conformation detected by quenched molecular dynamics (QMD; in a continuous dielectric of 80),[56,57] the sequential $C\alpha$ - $C\alpha$ distances were ~3.7 Å, whereas for a peptide in most ideal secondary structures are only slightly longer ~3.8 Å. This mimic has three chiral centers and no planar heterocyclic rings, so it is expected to extend chains in three dimensions that will vary amongst the eight (2^3) possible stereoisomers.

Using EKOS, preferred conformers of mimic **1** were systematically overlaid on common secondary structures. These ideal secondary structures are composed of L-amino acids, so it is tempting to assume that mimic LLL-**1** that has the same chain periodicity would tend to overlay better than the other stereoisomers, but this is *not* so (Figure 5.5a). The LLL-isomer is a better α- and 3_{10}-helical mimic than any other in the series, but select other stereoisomers overlay more closely on the other secondary structures. In fact, this is a trend observed throughout for all the chemotypes featured in this study.

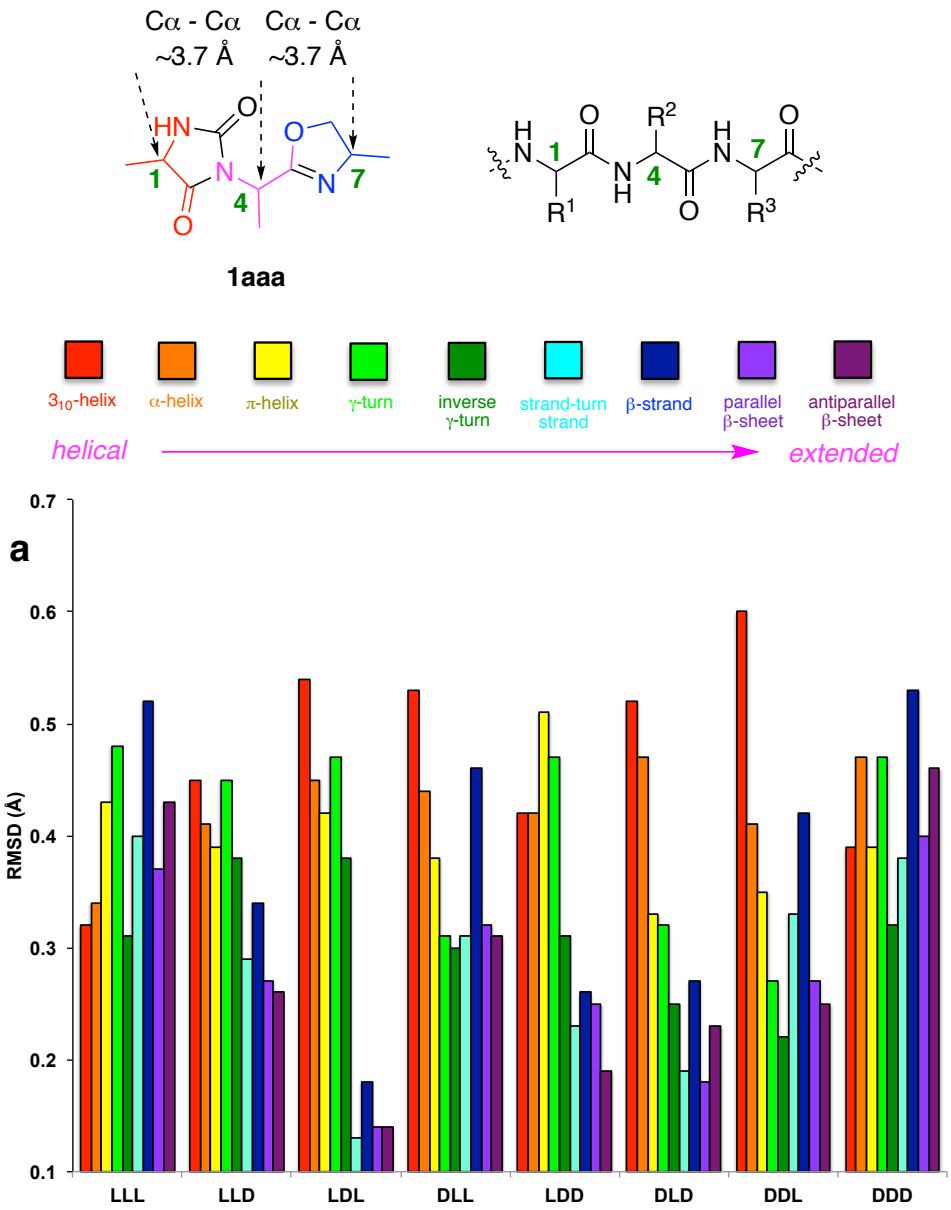
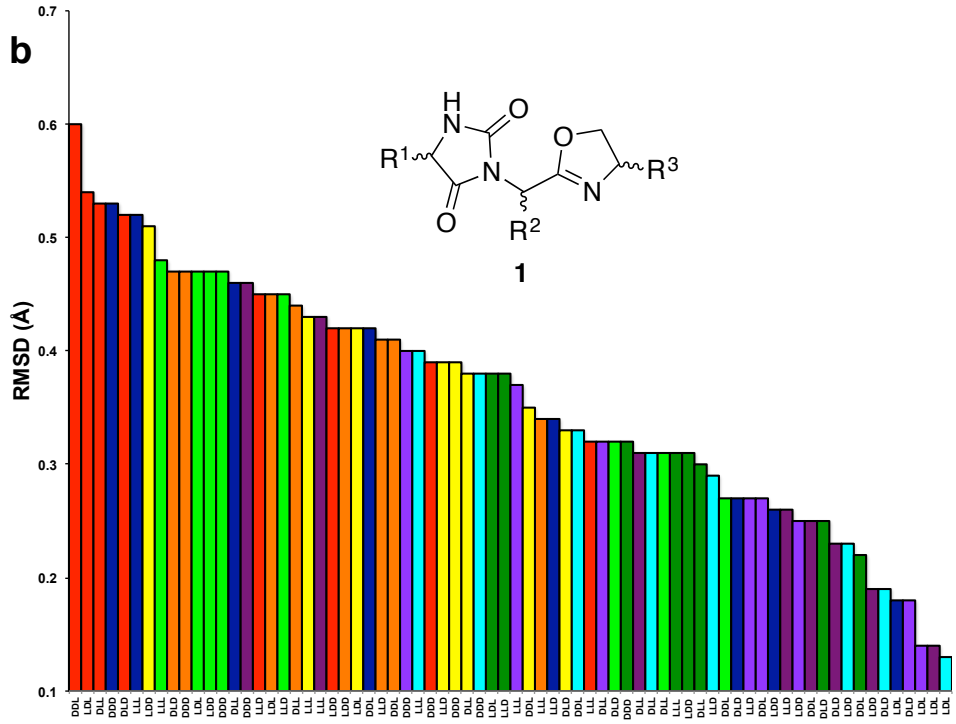


Figure 5.5 (**a**) RMSDs for overlay of the best matching accessible conformer of mimics **1** on each of the ideal secondary structures; (**b**) data in a replotted in descending RMSD (left to right) irrespective of stereochemistry. Statistical distribution of secondary structures at PPI interfaces derived by DSSP and STRIDE calculations; (**c**) the best 268 overlays of LLL-**1** (all RMSDs < 0.15 Å); and, (**d**) 1008 overlays of LDL-**1** (< 0.10 Å RMSD).



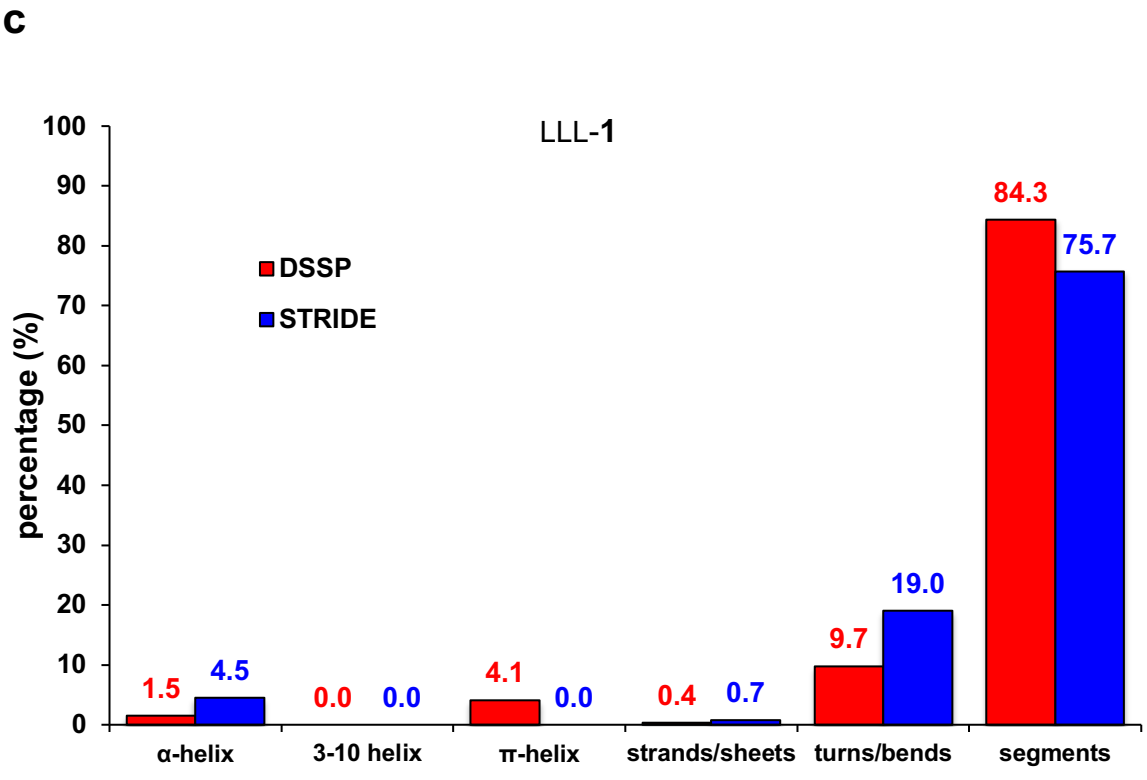


Figure 5.5 Continued.

d

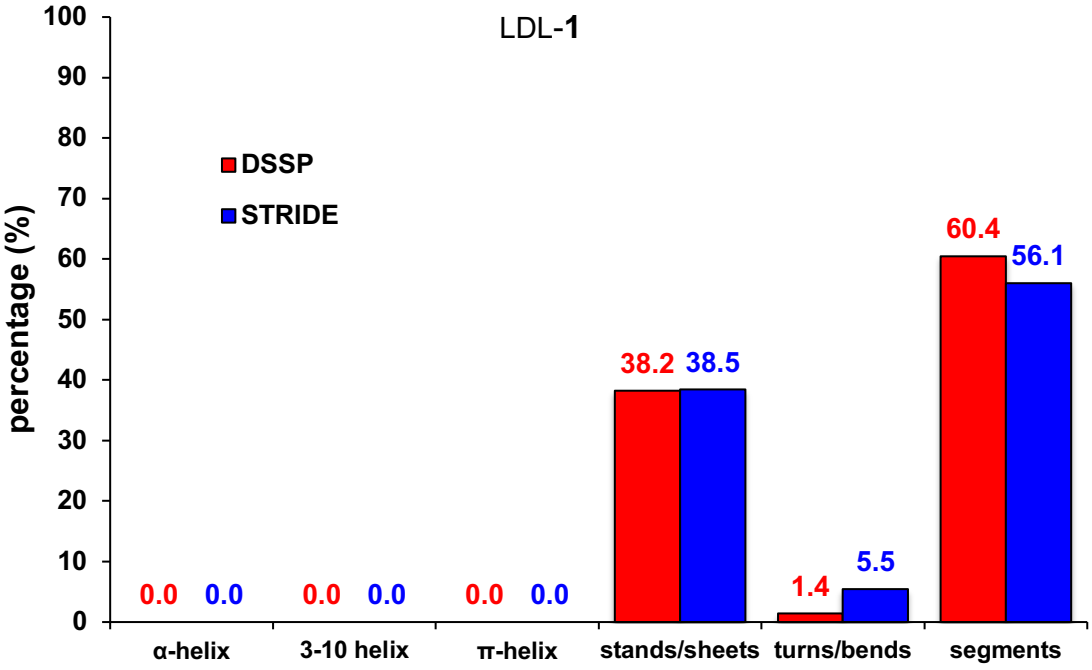


Figure 5.5 Continued.

Figure 5.5b replots the data in Figure 5.5a, but from highest to lowest RMSD, irrespective of secondary structure. This presentation reveals that compared with **A**, **1** is a superior mimic, especially for the extended conformations (0.13 - 0.18 Å RMSD): strand-turn-strand (light blue), β-strand (navy blue), parallel and antiparallel β-sheets (light and dark violet), *cf* blue and violet bars. Overlays of preferred conformers of **1** on more twisted helical structures (red, orange, yellow bars) occur at higher RMSDs. However, stereoisomers of chemotype **1** can be found to overlay on *any* of the ideal secondary structures with RMSD < 0.35 Å whereas, for comparison, chemotype **A** (Figure 5.4b; note the expansion of the y-axis showing RMSD, is different) only overlays well on antiparallel-, parallel β-sheets, and strand-turn-strand secondary structures with

RMSD < 0.35 Å. Relatively, chemotype **1** is therefore an outstanding universal mimic[53] of ideal secondary structures.

EKO mining of preferred conformers of LLL- and LDL-**1** on most of the PPI interfaces in the PDB showed that over 75 and 56 %, respectively, of the matches were *not* on ideal secondary structures at the interface (Figure 5.5c-d). The LDL-**1** stereoisomer is interesting insofar as it *does* match on sheet-type structures with a 38 % frequency; several DSSP and STRIDE analyses were performed for this study, but the data in Figure 5.5d is notable because it shows the highest bias among all nine chemotypes towards any secondary structure relative to "segments". Chemotype **1** was synthesized to validate its accessibility as described in Chapter II.

*5.2.3. Triazole-oxazole Chemotype **2**: No Difference Between Two Stereoisomers*

Chemotype **2** did *not* give preferred conformers that overlaid well with ideal secondary structures (< 0.35 Å). However, it is informative to consider the reasons why it is not such a good secondary structure mimic. Like the chemotypes **1**, the oxazole-triazole interface mimics **2** resemble continuous sets of three amino acids, *ie* tripeptide fragments. However, survey of favored conformers of **2** revealed their $C\alpha$ - $C\alpha$ distances tend to be around ~3.6 Å; this is shorter than for chemotype **1** and for tripeptides (~3.7 and 3.8 Å, respectively). Another significant difference between **1** and **2** is that the triazole-containing mimics **2** only have one chiral center. Flipping the stereochemistry of this sole chiral center does not afford the same degree of three-dimensional diversity as variation of three centers in chemotypes **1**. Consequently, it is unsurprising that none of the preferred conformers of **2** matched *any* of the ideal

secondary structures with RMSD < 0.35 Å (Figure 5.6b). Both stereoisomers of chemotype **2** were subjected to combine EKO and DSSP/STRIDE analyses. Both the L- and D-interface mimics of **2** showed a bias towards loosely wound helices (*ie* α- and π-) and towards turns, but not as dominant as uncategorized structure (Figure 5.6c-d). Compounds **2** were synthesized to explore how easily they could be made, as described in Chapter II. One compound in this series, L-**2fff**, crystallized and structure of this was evaluated via EKOX as shown in Figure 1.6 of Chapter I.

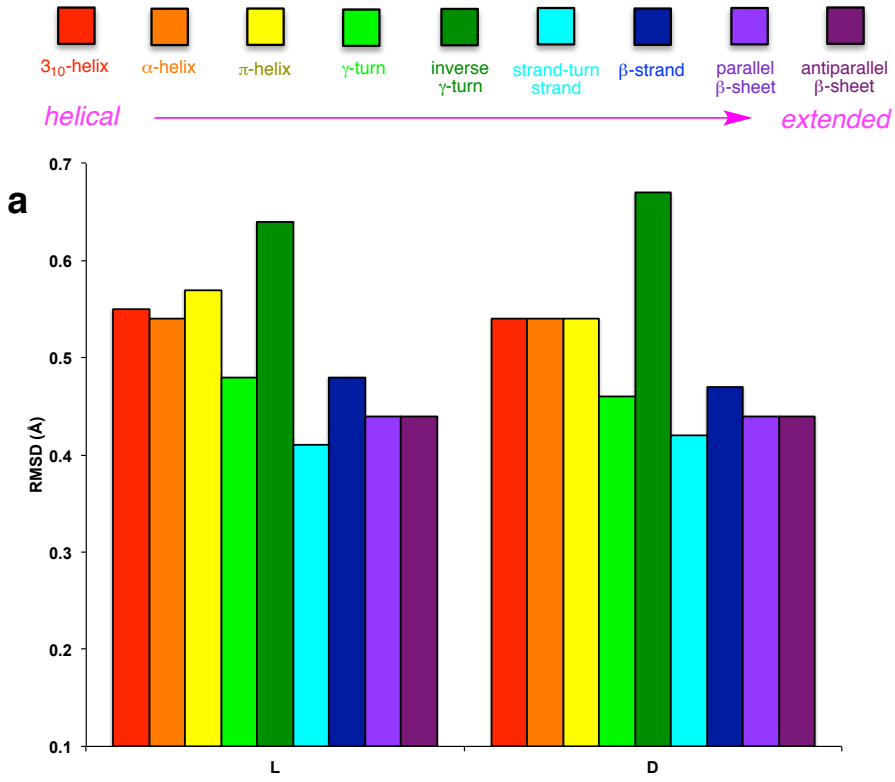


Figure 5.6 (**a**) RMSDs for overlay of the best matching accessible conformer of mimics **2** on each of the ideal secondary structures; (**b**) Replotted data in descending RMSD. Statistical distribution of secondary structures at PPI interfaces derived by DSSP and STRIDE calculations; (**c**) the best 106 overlays of L-**2** (all RMSDs <0.35 Å); and, (**d**) 98 overlays of D-**2** (<0.35 Å RMSD).

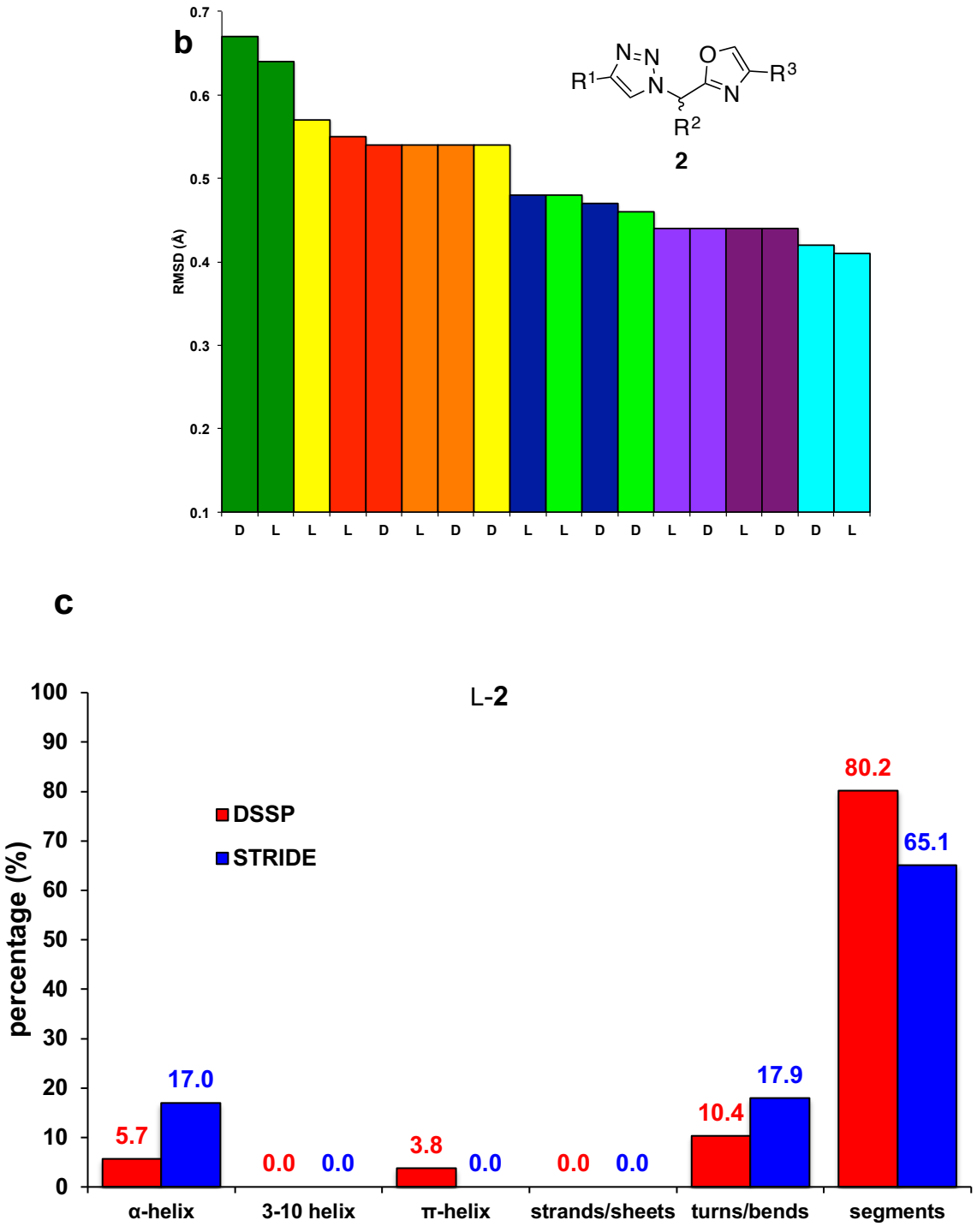


Figure 5.6 Continued.

d

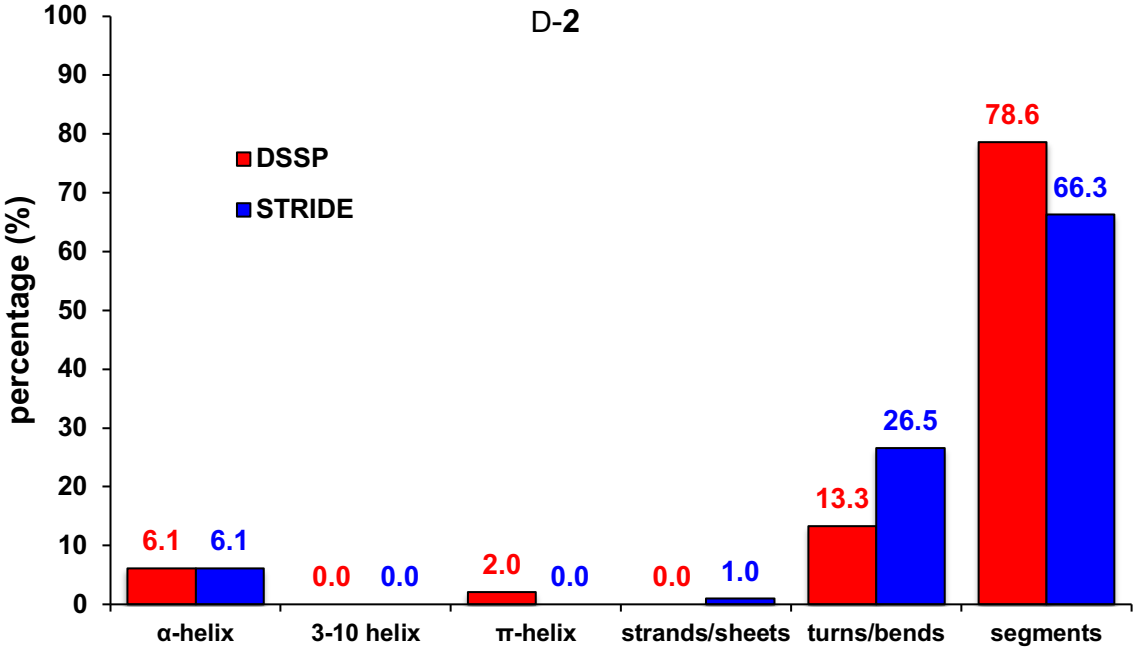


Figure 5.6 Continued.

*5.2.4. Bistriazole Oxazoline Chemotype **24**: A Surprisingly Helical Mimic*

Simulations in EKOS indicate mimic **24** can adopt a highly-populated conformation that overlays side-chains on the *i, i* + 4, *i* + 5 residues of a 3_{10}-helix and the *i, i* + 5, *i* + 6 residues of an ideal α-helix. This observation is unlike those in analyses of **A**, **1** and **2** and, as it transpires, of **25** - **29**; none of those chemotypes gave *any* helical low energy conformers that overlaid with RMSDs < 0.30 Å (see Figure 5.5a for the second best helical mimic LLL-**1**, on a 3_{10}-helix, 0.32 Å). Collectively, these data indicate **24** is the best helical mimic in the series. Further consideration indicates helical conformations are favored by mimics **24** with an L-configuration at the second position; the only exception is for DDD-**24** which the simulations indicate also adopts helical conformations.

Conversely, mimics **24** with a D-configuration at the second position tend to adopt extended conformations (again, except for DDD-**24**). Overall, by a small margin, the best overlays observed for **24** in EKOS analyses were for anti- and parallel-β-sheet motifs (*ca* 0.20 and 0.21 Å, respectively; Figure 5.7b). The fact that none of the preferred conformers of **24** overlaid on γ-turns and β-strands is logical; the mimic is too long to fit γ-turns and too kinked to overlay well on the linear β-strands motif.

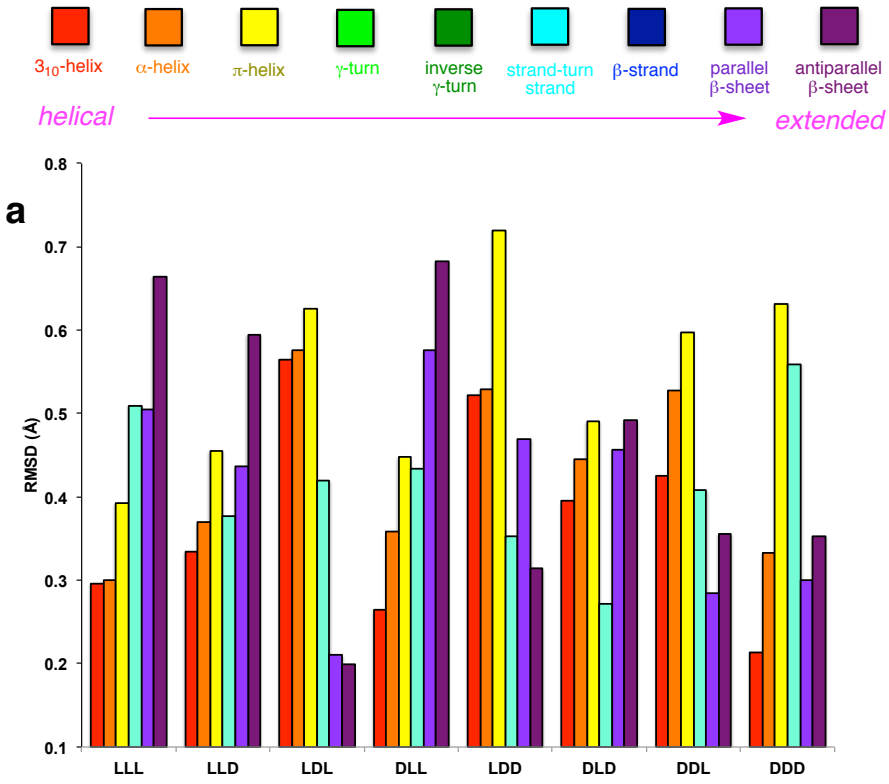
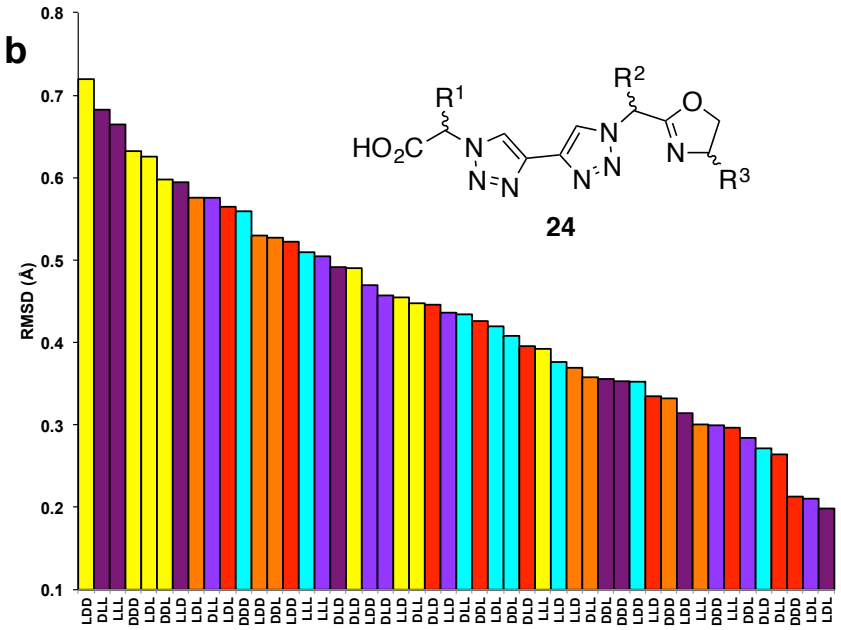


Figure 5.7 (**a**) RMSDs for overlay of the best matching accessible conformer of mimics **24** on each of the ideal secondary structures; (**b**) Replotted data in descending RMSD. Statistical distribution of secondary structures at PPI interfaces derived by DSSP and STRIDE calculations; (**c**) the best 308 overlays of LLL-**24** (RMSDs < 0.30 Å); and, (**d**) 369 overlays of LDL-**24** (RMSDs < 0.25 Å).



c

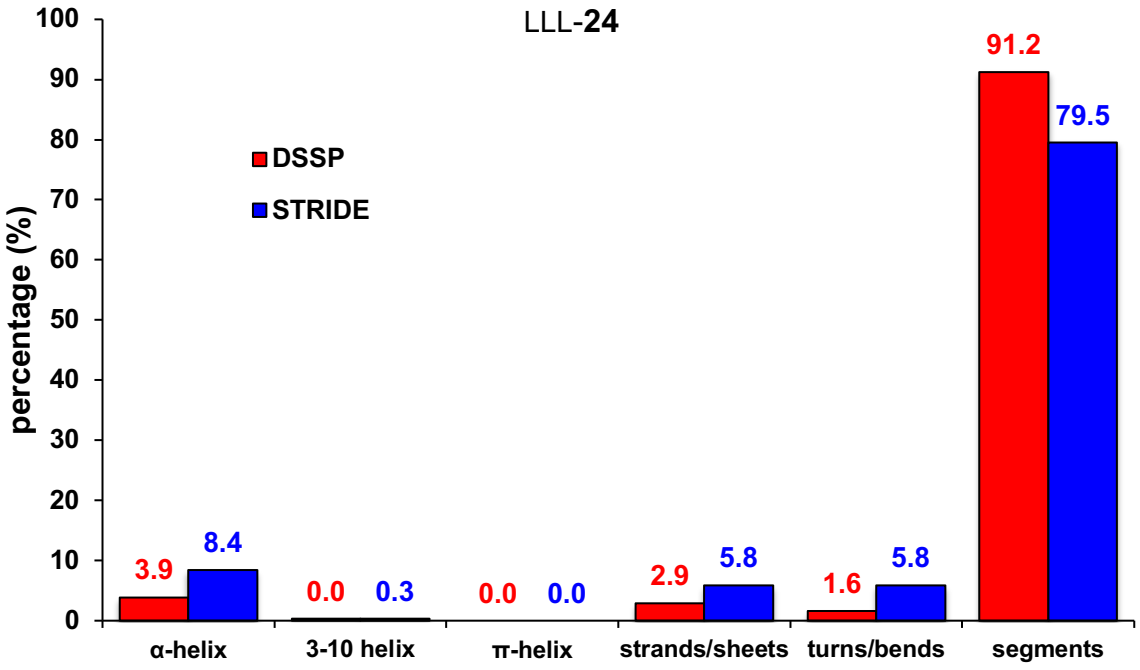


Figure 5.7 Continued.

d

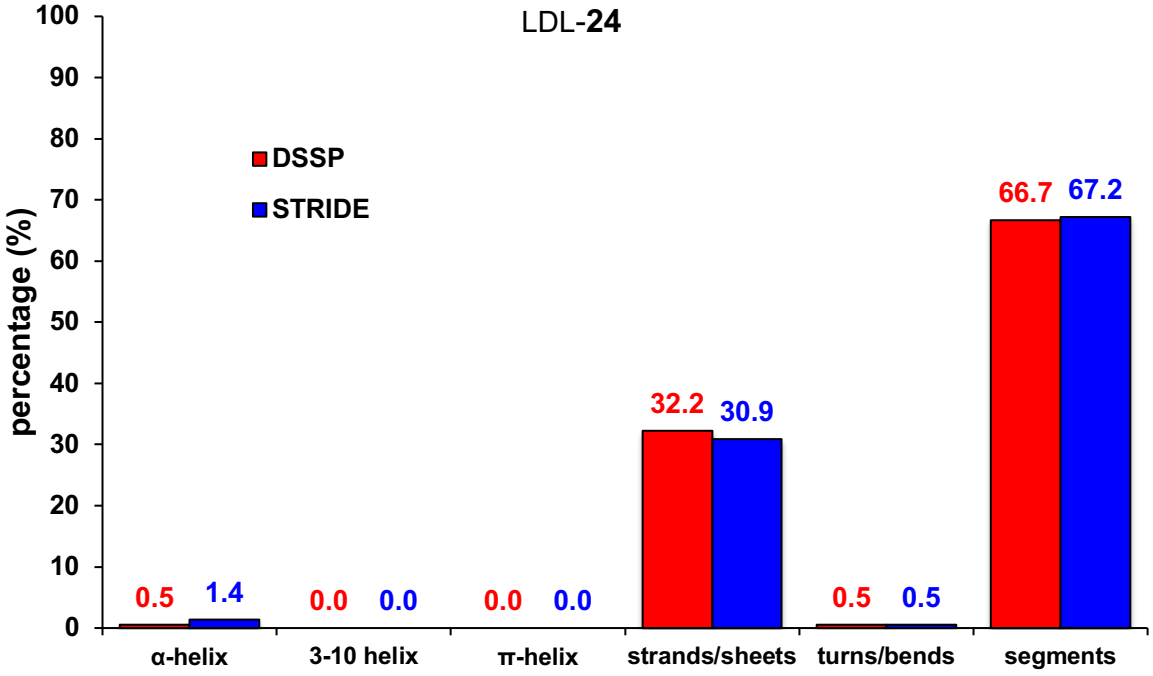


Figure 5.7 Continued.

In EKO analyses featuring PPI interfaces, LLL-**24** overlaid most of the time on "segments" but showed a small bias towards α-helices out of all secondary structures. By contrast, after segments, LDL-**24** shows a bias towards sheets at PPI interfaces (Figure 5.7c-d). These two observations are consistent with the conclusions from the EKOS analyses.

*5.2.5. Dihydantoin Chemotype **25**: Short Side-chain Separation*

It now appears that chemotypes **25** are good-to-poor interface mimics, based on several observations. First, the R^1 and R^2 on hydantoin motif are quite close together relative to dipeptides, so preferred conformers of this chemotype did not overlay on any ideal

secondary structure with RMSD <0.35 Å in EKOS analyses. Above that threshold, loose overlays were observed for strand-turn-strand and α-helices (*i*, *i*+1, *i*+9 side-chains). When matched with PPI interfaces (EKO), the second frequent overlays apart from segments were on a turns/bends region classified by DSSP/STRIDE.

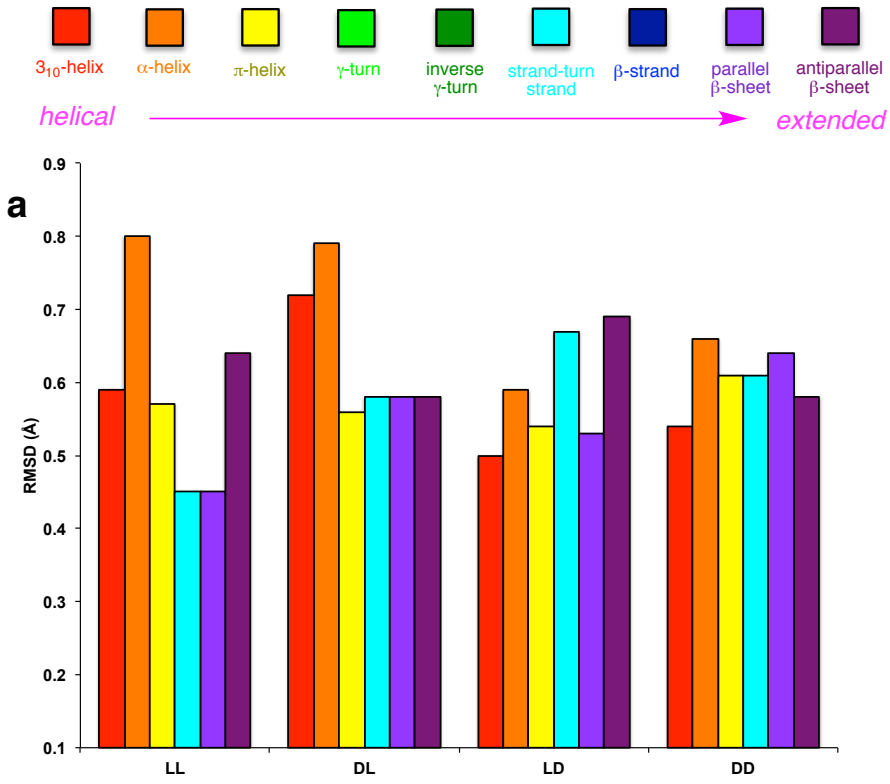
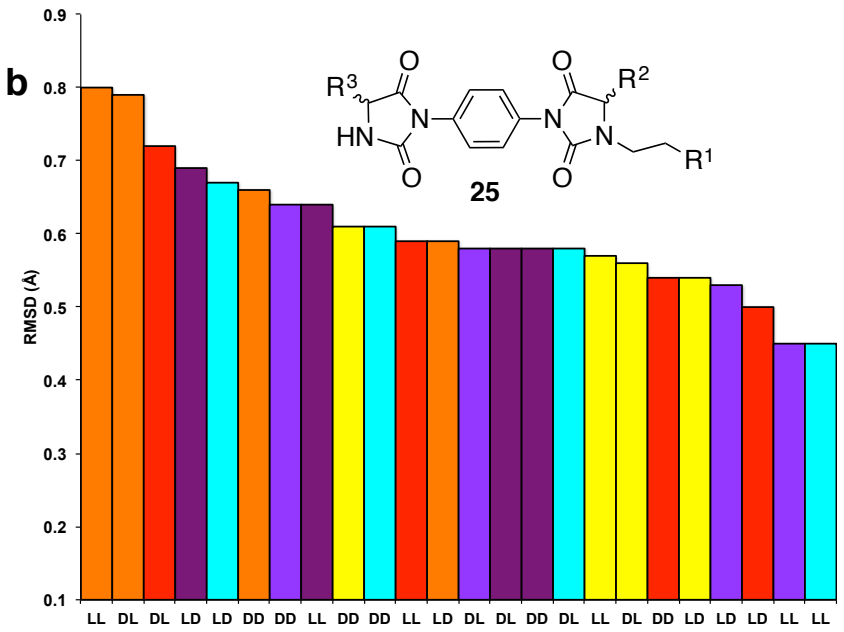


Figure 5.8 (**a**) RMSDs for overlay of the best matching accessible conformer of mimics **25** on each of the ideal secondary structures; (**b**) Replotted data in descending RMSD. (**c**) Statistical distribution of secondary structures at PPI interfaces derived by DSSP and STRIDE calculations of the best 120 overlays of LL-**25** (<0.35 Å RMSD) on PPI interfaces



c

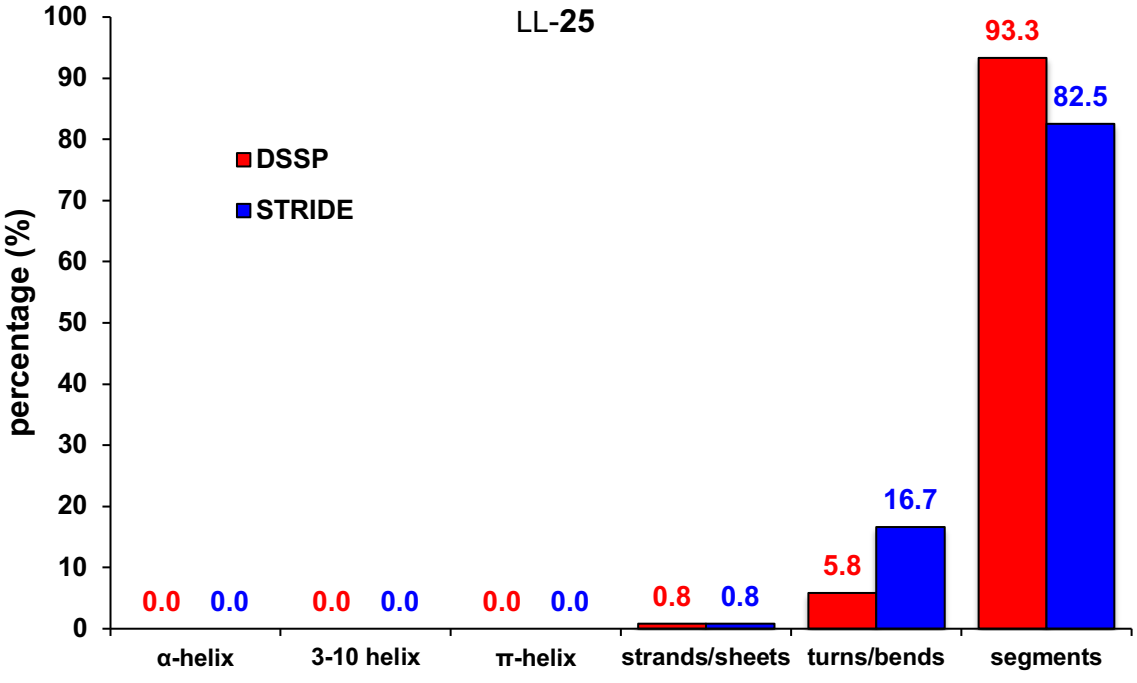


Figure 5.8 Continued.

*5.2.6. Hydantoin Oxadiazole Oxazoline Chemotype **26**: A Good Extended-sheet Mimic*

Generally, mimic **26** gives an excellent match on sheets, some good matches on turns, and barely hits for helices (Figure 5.9). The "periodicity" of the hydantoin fragment that projects the R^1 and R^2 substituents of this mimic is identical to consecutive residues in peptides, *ie* both contain $NCHR^1CONCHR^2CO$ motifs, (*cf* mimic **1**). In fact, the *C*α - *C*α seperation in the hydantoin fragment (3.8 Å) almost perfectly matches of that in an extended peptide conformation. The R^3 side-chain is further away (~6.8 Å), so to overlay with secondary structures this substituent must align with a distal residue. If this occurs, it seems reasonable that chemotype **26** would be a good mimic of sheets.

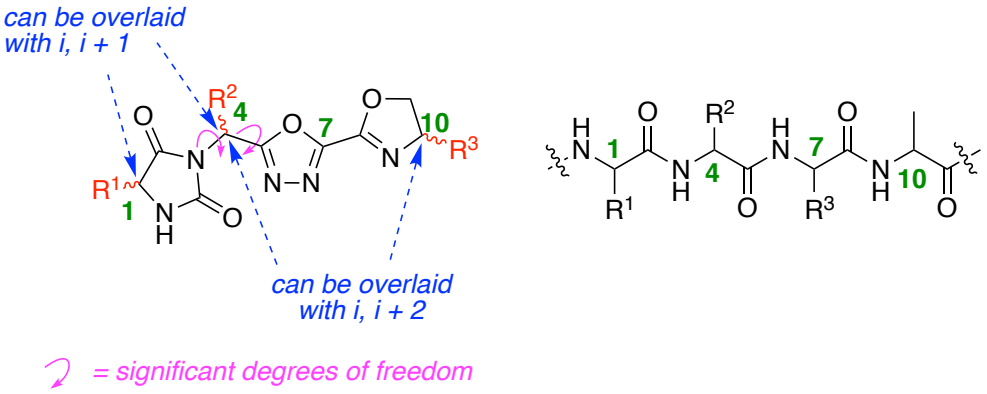


Figure 5.9 (**a**) RMSDs for overlay of the best matching accessible conformer of mimics **26** on each of the ideal secondary structures; (**b**) Replotted data in descending RMSD. Statistical distribution of secondary structures at PPI interfaces derived by DSSP and STRIDE calculations; (**c**) the best 287 overlays of LDD-**26** (all RMSDs <0.18 Å); and, (**d**) 115 overlays of DDD-**26** (<0.20 Å RMSD).

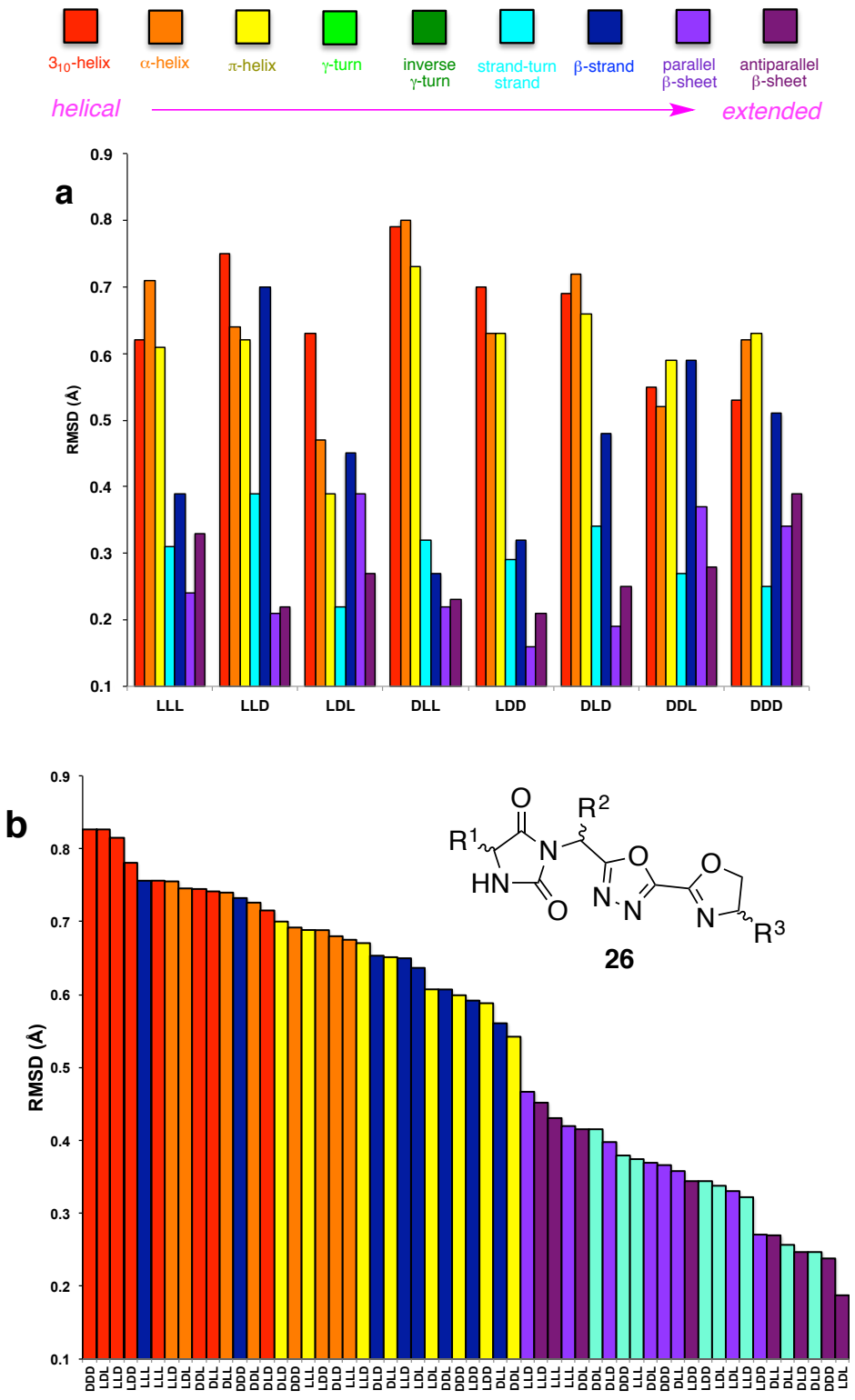
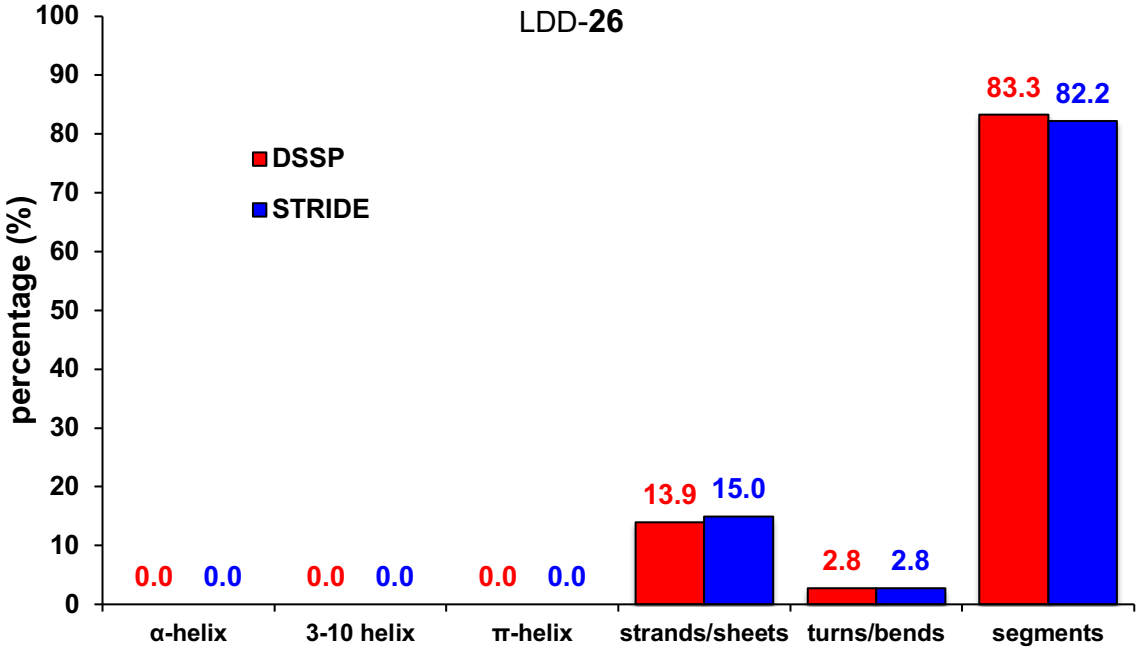


Figure 5.9 Continued.

c



d

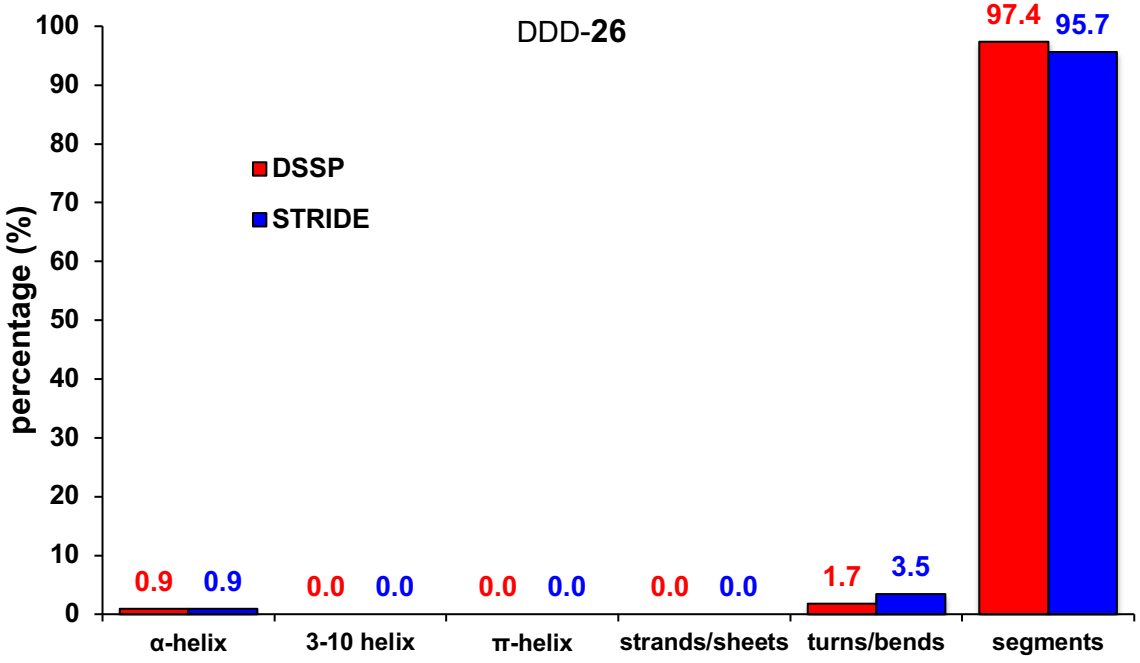


Figure 5.9 Continued.

The $C\alpha$ - $C\alpha$ seperation for *i* and *i + 1* residues in ideal antiparallel β-sheets is 3.9 Å, *ie* slightly longer than in parallel sheets (~3.8 Å), while the separation of *i* and *i + 2* residues is slightly less at 6.6 Å (as opposed to 6.8 Å). As a result, matching of mimic **26** on an ideal antiparallel sheet is not as good as on parallel sheets, but still acceptable. Mimic **26** also overlays on the *i*, *i + 1*, and *i + 5* residues of an ideal π-helix, but with RMSD values > 0.35 Å. The reason why this chemotype does not overlay very well is that the *i + 1* to *i + 5* distance in an ideal π-helix is only 6.3 Å, but mimic **26** projects the third side-chain at more distal positions (around 6.8 Å). Similarly, that long R^2 to R^3 separation makes it impossible for **26** to overlay on turns.

Mimic **26** was compared with a huge number of PPI interfaces using the EKO routine. Approximately 300 best-fitting conformers (all RMSD < 0.18 Å) were selected for the stereoisomer that best fit secondary structures in general, *ie* the LDD-form. Most of the selected conformers matched on "segments" in the DSSP/STRIDE analysis, but approximately 15 % matched on sheets. To test the relevance of matching on secondary structures, we also consider the DDD-isomer because it gave the *worst* matches in the stereoisomeric series. In the EKO analysis of PPI interfaces, it was necessary to raise the RMSD threshold slightly, to 0.20 Å, to obtain 116 matches; > 95 % of those matched on "segments" (Figure 5.9c-d). These observations are consistent with the data from the EKOS analyses; the LDD-isomer tends to match on sheets more than other secondary structures, while the DDD-isomer has no obvious preference for any secondary structures.

5.2.7. *Imidazolidinone Bisoxazoline Chemotype* ***27****: Another Sheet Mimic*

Chemotypes **26** and **27** are similar insofar as they have two rigid fragments linked by two rotatable bonds. In fact, both cores consist of three, 5-membered heterocyclic rings, wherein the ones that support the R^1 and R^2 side-chains are seperated by two rotatable bonds. Consequently it is unsurprising that, like **26**, mimic **27** shows a bias towards sheet-like conformations. However, the fit is not quite as good: the *C*α - *C*α separations (R^1 to R^2, 3.7 Å; R^2 to R^3, 7.2 Å) are a little too long.

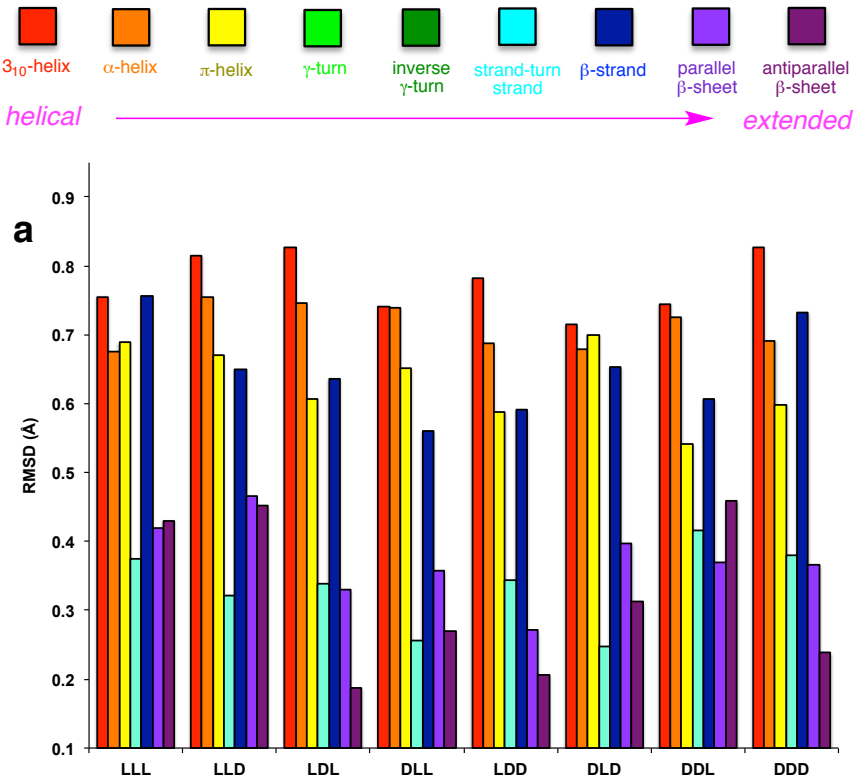
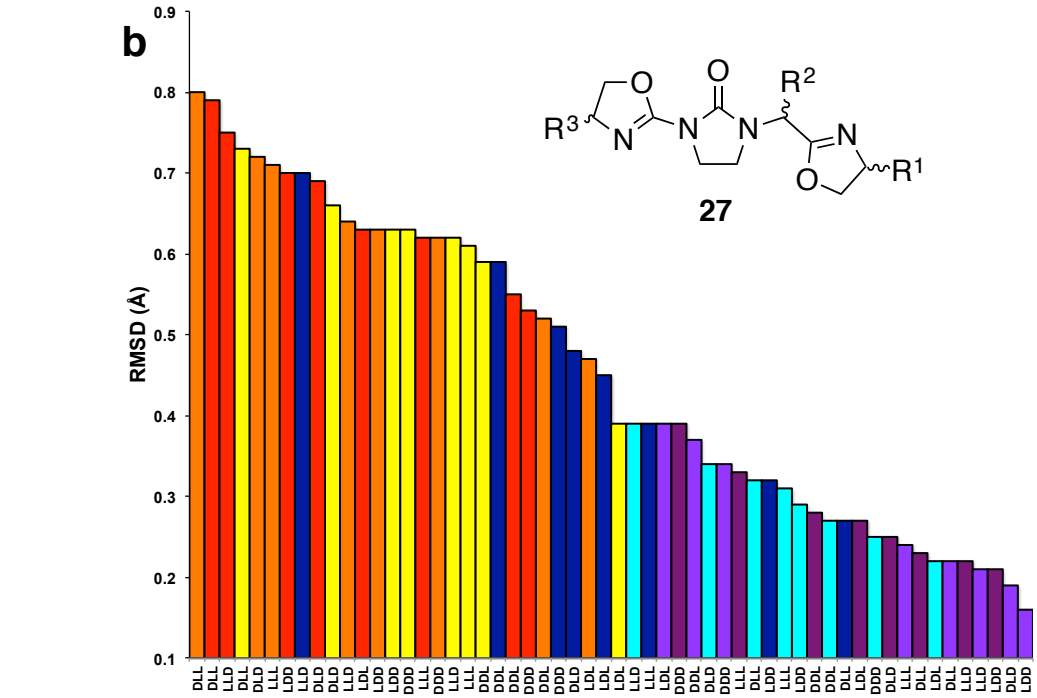


Figure 5.10 (**a**) RMSDs for overlay of the best matching accessible conformer of mimics **27** on each of the ideal secondary structures; (**b**) Replotted data in descending RMSD. (**c**) Statistical distribution of secondary structures at PPI interfaces derived by DSSP and STRIDE calculations of the best 288 overlays of LDL-**27** (all RMSDs <0.20 Å)



c

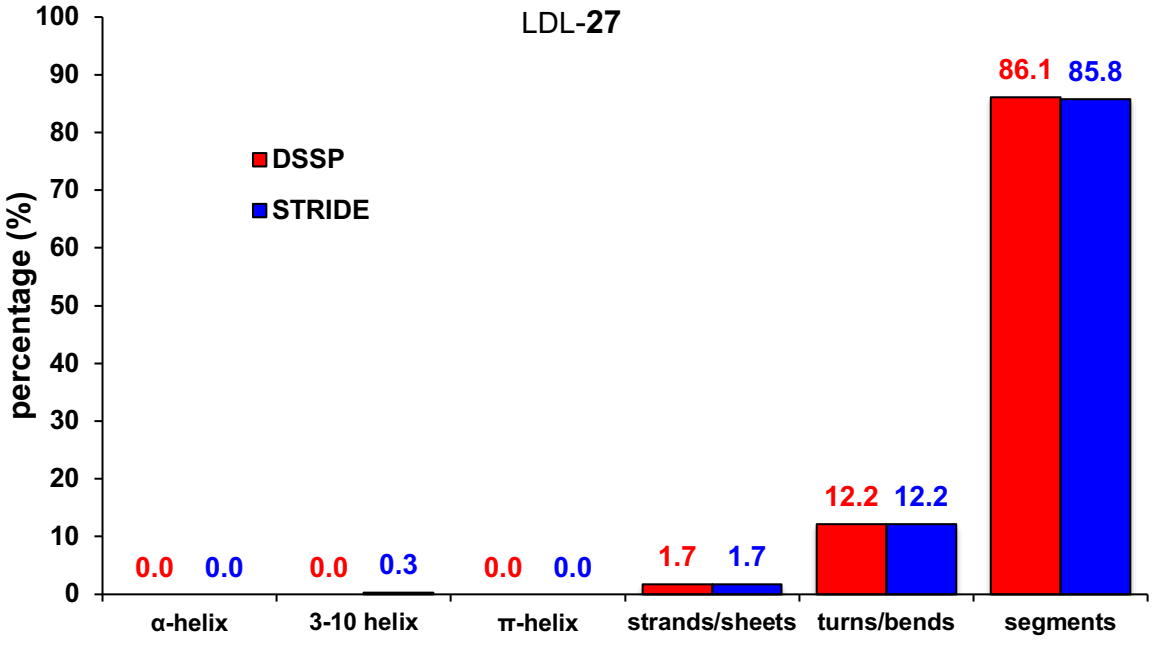


Figure 5.10 Continued.

5.2.8. *Bistriazole Chemotype* ***28****: Long Side-chain Separation*

Triazole motifs in mimic **28** rigidly hold the *C*α – *C*α atoms 4.9 Å apart, and this is longer than the corresponding separation in a typical peptide (~3.8 Å). This single observation accounts for the fact that chemotypes **28** did not overlay on *any* ideal secondary structures with RMSDs of <0.35 Å. Within that group of poor matches, this structure overlays best on sheets, then 3_{10}-helices, but not at all well on strands and π-helices (Figure 5.11b).

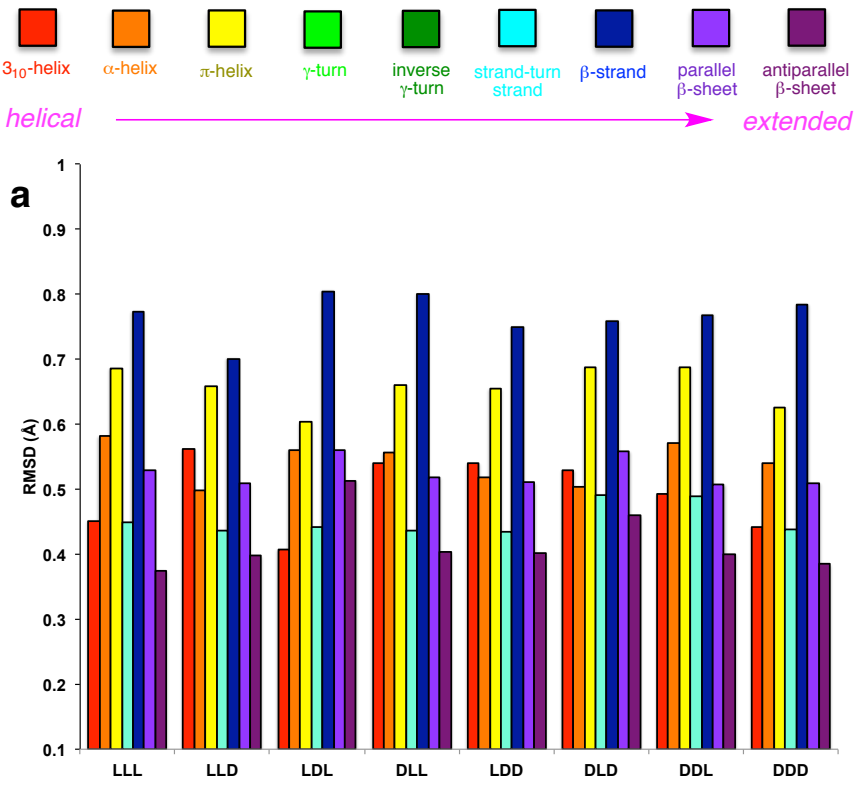
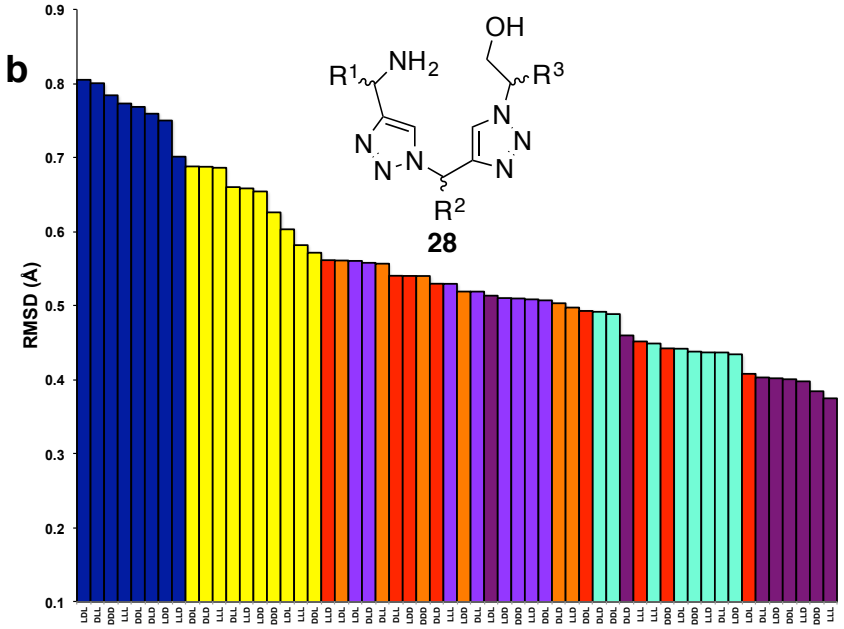


Figure 5.11 (**a**) RMSDs for overlay of the best matching accessible conformer of mimics **28** on each of the ideal secondary structures; (**b**) Replotted data in descending RMSD. (**c**) Statistical distribution of secondary structures at PPI interfaces derived by DSSP and STRIDE calculations of the best 181 overlays of LLL-**28** (<0.50 Å RMSD) on PPI interfaces.



c

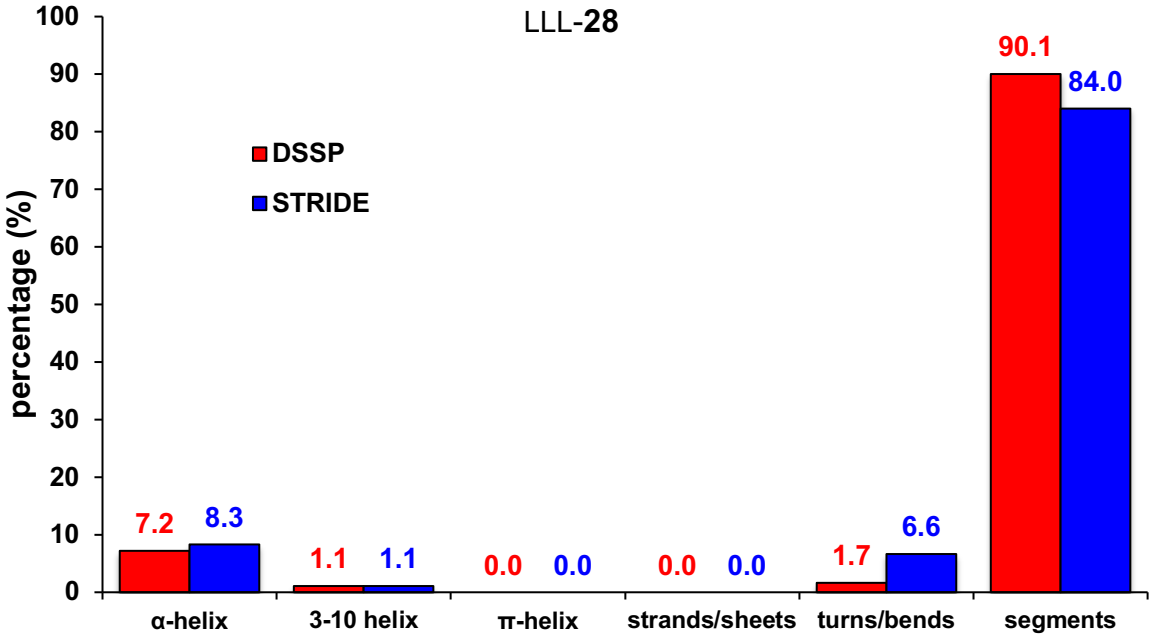


Figure 5.11 Continued.

*5.2.9. Compact Bistriazole Chemotype **29**: Poor Secondary Structure Mimicry*

Preferred conformations of chemotype **29** gave even worse overlays on ideal secondary structures than **28**. This mimic has more appropriate $C\alpha$ – $C\alpha$ separations for peptide mimicry, but the orientations of the $C\alpha$ – $C\beta$ vectors are locked into the same plane as the triazole rings. Interestingly, there are significant different results between DSSP and STRIDE on the turns/bends motif (Figure 5.12c). This is likely due to the flexible criteria of STRIDE that allow more deviation than DSSP.

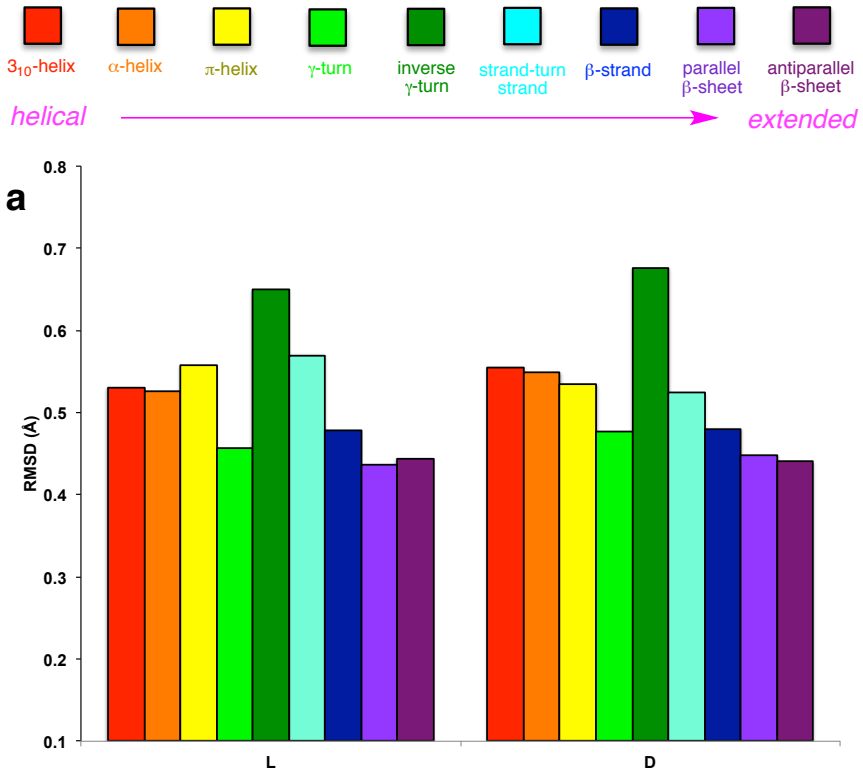
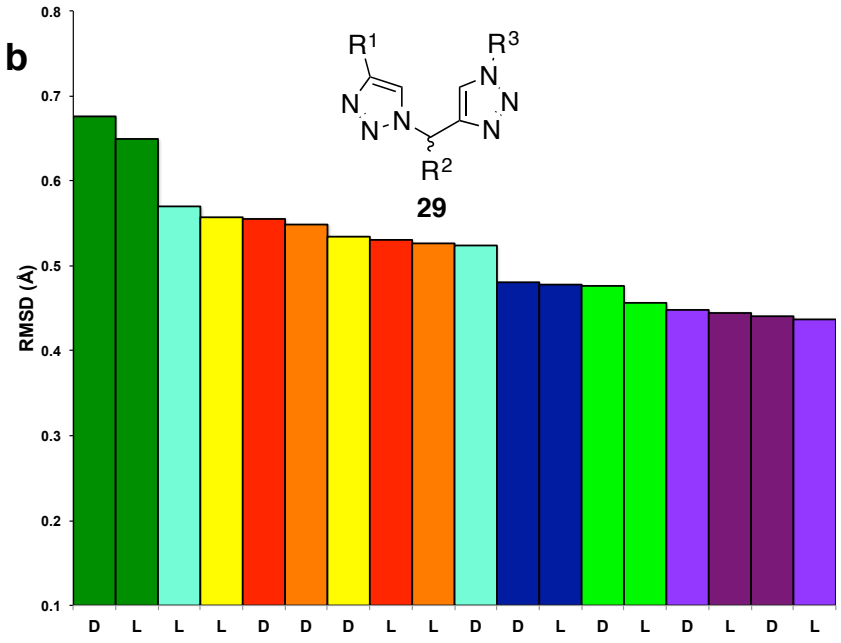


Figure 5.12 (**a**) RMSDs for overlay of the best matching accessible conformer of mimics **29** on each of the ideal secondary structures; (**b**) Replotted data in descending RMSD. (**c**) Statistical distribution of secondary structures at PPI interfaces derived by DSSP and STRIDE calculations of the best 654 overlays of L-**29** (<0.50 Å RMSD) on PPI interfaces.



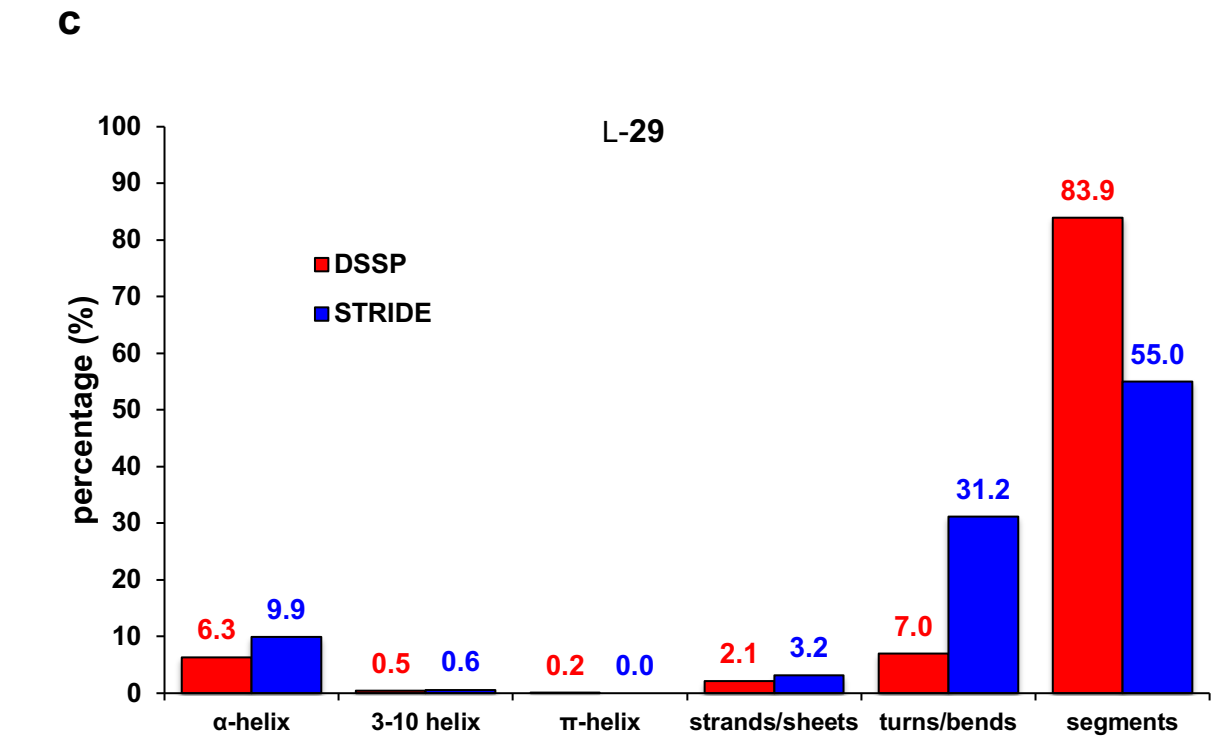
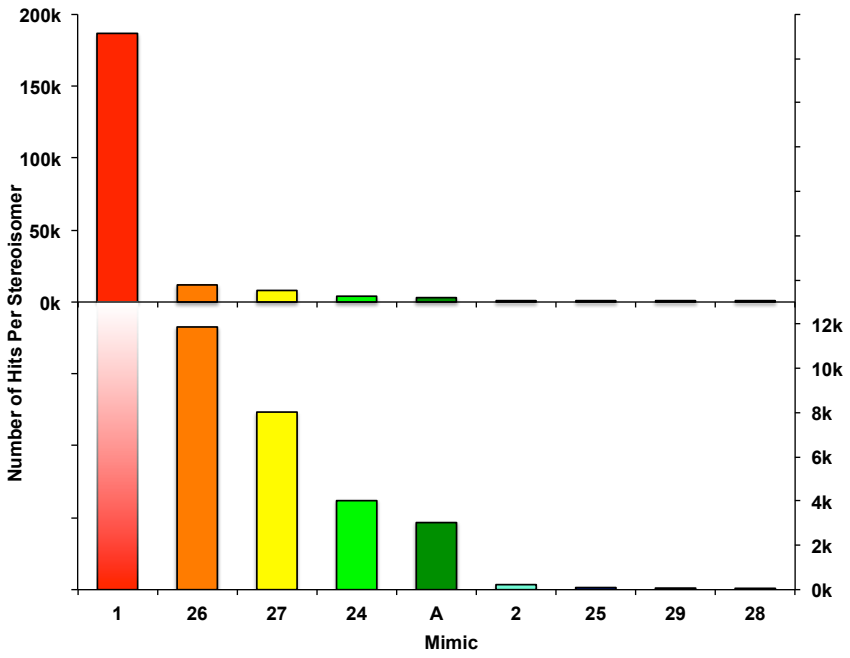


Figure 5.12 Continued.

5.3 Conclusion

Aims of this work were to: (i) determine how these minimalist mimics compare with huge numbers of interfaces in crystallographically characterized PPIs; (ii) compare this data with the bias of the same mimics to adopt conformations that mimic secondary structures; and, (iii) evaluate the presence of secondary structures at the interfaces where the mimics do overlay well. To summarize our findings towards the first aim, Figure 5.13a plots the number of EKO hits for each chemotype (all stereoisomers combined) with RMSDs <0.35 Å. Chemotype **1** had 1,491,780 overlays, whereas the second-best interface mimic (**26**) had only 94,972. Mimics **1**, **26**, **27**, and **24** gave more matches than our "control" **A**, whereas **2**, **25**, **29**, and **28** matched less frequently. Thus, **26**, **27**, **24** and **A** are fine interface mimics, and chemotype **1** seems to be a truly privileged structure. One reason for this is that **1** incorporates one of the best dipeptide fragments analyzed in this work (see explanation in Appendix E).

a



b

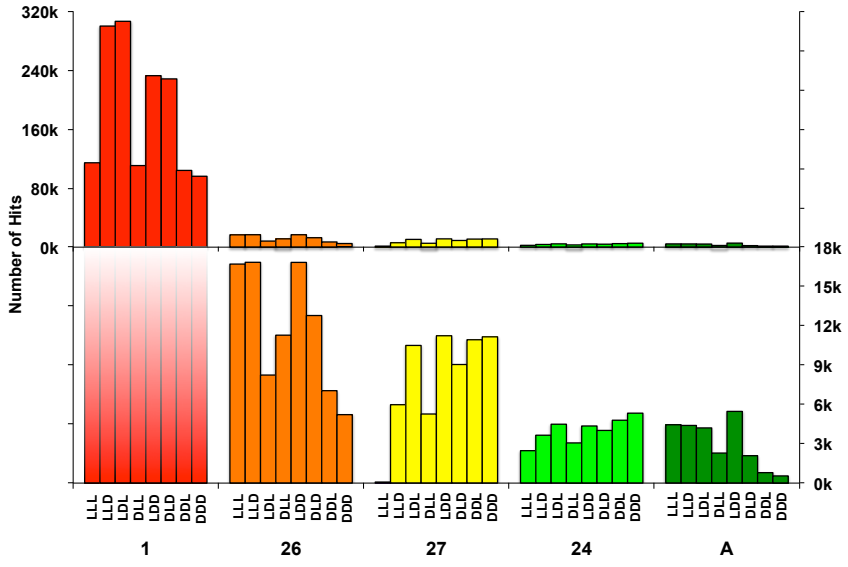


Figure 5.13 Number of EKO hits with RMSD < 0.35 Å: (**a**) totals for each chemotype irrespective of stereochemistry for each chemotype; and, (**b**) per stereoisomer for **1**, **26**, **27**, **24** and **A**.

Our second aim was to compare this data with the bias of the same mimics to adopt conformations that mimic secondary structures. Consideration of the EKOS data in each chemotype reveals that the potential of chemotypes **1**, **26**, **27**, **24** and **A** for secondary structure mimicry ranks in exactly that order and this *corresponds to the relative number of EKO hits (**1** >> **26** > **27** > **24** > **A**) from Figure 5.13.* Chemotypes **24**, **26**, **27** and **A** are superior minimalist mimics of secondary structures, and **1** is truly exceptional. Chemotype **1** is also remarkable because minimalist mimics that display three side-chains in orientations that align with three *consecutive* peptidic side-chains are rare; in fact, there may be only one well-known example, the Smith-Hirschmann β-strand mimics.[40,185]

The third aim of this study was to evaluate the presence of secondary structures at the interfaces where the mimics do overlay well. If all three side-chains matched in the DSSP or STRIDE analyses overlaid on an interface secondary structure then that overlay was "binned" accordingly, but if one of them did not then the overlay was recorded as being on a "segment". Segment overlays predominated for every stereoisomer on each of the chemotypes examined, without exception, and there were only a few instances for which a bias towards any secondary structure represents over 30 % of the top hits. Chemotype LDL-**1** was exceptional because it matched sheets 38 % of the time, and LDL-**24** matched the same secondary structures in 31 % of the cases. Mimic L-**29** matched turns/bends more than 30 % only in the STRIDE analysis, which was arguable about the bias towards these structures. By comparison, mimic LDD-**A** had the next highest frequency, matching sheets with a 15 % hit-rate. The overall, conclusion from analysis of >240,000 interfaces is clear: *the particular preferred conformers of the minimalist mimics that overlaid well on PPI interfaces did not tend to do so on secondary structure interface motifs; they overlaid far more often on interface*

regions that only partially overlap with a secondary structure motif, or do not overlap with one at all.

To summarize, the overwhelming trend in the data collected indicates **1**, **26**, **27**, **24** and **A** overlaid well on secondary structures and on many PPI interfaces. However, in cases where they did overlay on protein-protein interfaces, it was most often *not* on secondary structure motifs. Thus, this study comes to a pivotal conclusion: *secondary structure mimicry is only an indirect indicator of interface mimicry*.

CHAPTER VI

PROTEIN-PROTEIN INTERFACE MIMICRY BY AN OXAZOLINE PIPERIDINE-2,4-DIONE*

6.1 Introduction

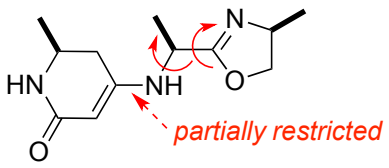
Early minimalist mimics of secondary structures inspired by Hamilton's terphenyls[27-29,41,186,187] featured planar aromatic units that display side-chains in appropriate orientations. More recently this field has turned toward chiral and heterocyclic designs that exist in coiled conformations, and/or tend to have superior water solubilities.[188] It is convenient if mimics of this kind can be produced from amino acid starting materials, otherwise it is difficult to incorporate all the different side-chains. The recently reported oligooxopiperazines, for example, are derived from amino acids.[38]

Even though effective minimalist mimics are more rigid than peptides,[52] most populate multiple solution conformers that display side-chains in different orientations.[53] We developed two strategies, *EKO*[58] and *EKOS*,[54] to ascertain how conformations of minimalist mimics resemble protein-protein interface regions and ideal secondary structures, respectively. Application of EKO exposes the enormous diversity of PPI interfaces: even a small fraction of these could not be accurately represented by all the secondary structure mimics reported in the literature to date. Consequently, there is a need to develop and understand new chemotypes for the key issue of interface mimicry.

---

* Reprinted with permission from "Protein-Protein Interface Mimicry by an Oxazoline Piperidine-2,4-dione", Li, X.; Taechalertpaisarn, J.; Xin, D.; Burgess, K. *Org. Lett.*, **2015**, *17*(3), 632-635. Copyright 2017 American Chemical Society.

This research introduces chiral, nonaromatic, interface mimics **30** composed of piperidine-2,4-dione and oxazoline fragments linked by −NHCHR− units (Figure 6.1). The objectives of this study were to develop a synthesis of molecules **30** that could incorporate many genetically encoded amino acid side chains, to elucidate the bias of this scaffold toward all the common ideal secondary structures and to show illustrative cases where EKO predicts an excellent match of accessible conformers of **30** on PPI interface regions.



L,L,L-**30aaa**

MM 237, no amide bonds

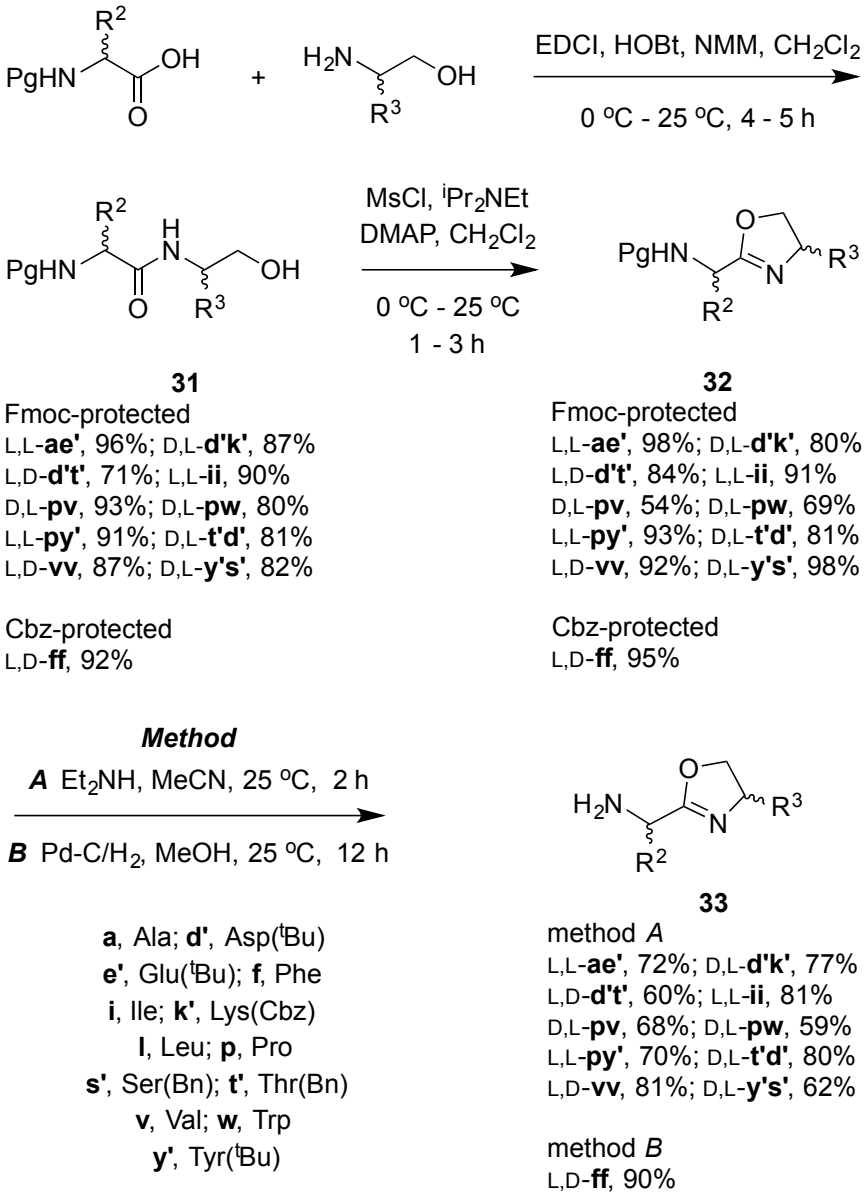
predicted properties:
pCaco 1118 nm/s
log o/w partition coefficient 1.54
polar surface area 76.97
no rule of 5 violations

Figure 6.1 Scaffold **30** is a minimalist mimic of secondary structures with favorable predicted properties for cell and oral bioavailability.

6.2 Synthesis of Oxazoline Piperidine-2,4-dione

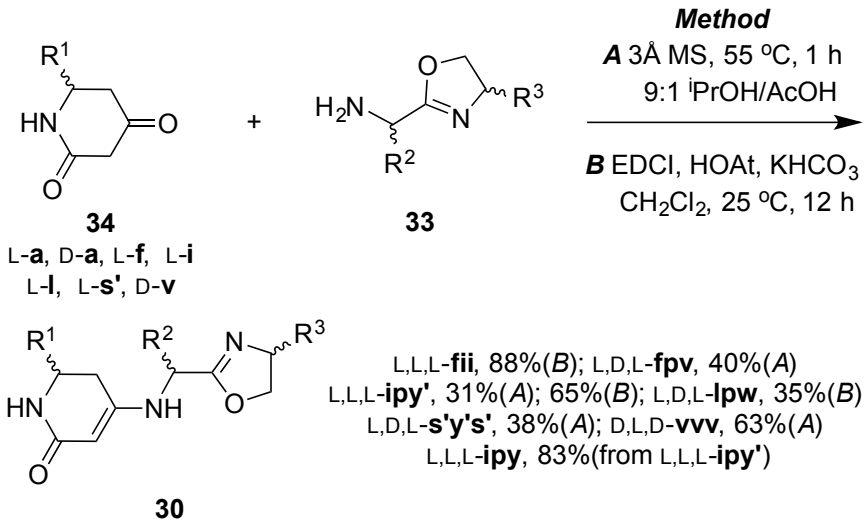
Scheme 6.1 shows how the oxazoline fragments were prepared from Fmoc-protected or Cbz-protected amino acids and amino alcohols. After a routine coupling to obtain molecules **31**, the primary alcohol was mesylated and then treated with base to initiate oxazoline formation. Some Fmoc-protected compounds related to **32** have been reported prior to this work,[110,189] but most of the systems with the side chains indicated in Scheme 6.1 have not been prepared before. The cyclization conditions in Scheme 6.1 were arrived at after some optimization; they are a modification of those used in Sigman's aminooxazoline syntheses.[110] Many other conditions that did not use DMAP or relied upon activation via PPh_3/CCl_4 gave poor product yields. Removal of the Fmoc

protecting group from the protected amines **32** gave the aminooxazolines **33**. A similar procedure was used, but with *N*-Cbz protected Phe, to access the **ff** chiron.



Scheme 6.1 A Fmoc approach was used to obtain most chirons **33**, but Cbz was used to access the **ff** chiron.

Having obtained a set of aminooxazolines, we developed two methods (A and B[190] in Scheme 6.2) to add these to the piperidine-2,4-dione derivatives **34**;[55,191,192] product was obtained using either approach, but the yields differed on a case-by-case basis. Overall, the synthetic route is divergent−convergent because any ketone **34** can be condensed with any amine **33**.



Scheme 6.2 Syntheses of the target compounds **30**.

6.3 Evaluation of Simulated Conformers By EKOS

EKOS was used to relate the ensemble of simulated accessible conformers of **30aaa** to ideal secondary structures. This process was carried out for all stereoisomers of **30aaa**, full data are shown in the Appendix F, and two select examples are given here. Data here were calibrated relative to the average RMSDs for all the best fitting conformers for a given stereomer. Thus, for LLL-**30aaa**, 0.41 Å was the average for the seven best fitting conformers on the seven secondary structures indicated. Conformers LLL-**30aaa** matched better on an ideal α-helix than on any of the other secondary structures (Figure

6.2a). In our experience, it is much harder to design good helical mimics than ones that overlay other motifs. Our application of EKOS on oligooxopiperazines **A**[38] (all stereoisomers, Chapter V) indicate they have accessible conformations that overlay well on an ideal α-helix with a RMSD of 0.46 Å (DLL-**A,** based on the 6 *C*α and *C*β coordinates), and that is better than nearly all of the other mimics of ideal α-helical conformations in the literature.[54] However, LLL-**30aaa** appears to be a superior α-helical mimic since it can adopt a conformation that matches an ideal α-helix with a RMSD of only 0.26 Å.[54]

Achiral minimalist mimics like terphenyls have only one isomer to compare with ideal secondary structures. Conformations of *chiral* minimalist mimics, however, are stereochemically dependent, and we offer two observations related to this. First, stereochemical changes can *significantly* alter the conformational bias of many minimalist mimics such that one isomer can match extended conformations whereas another is more closely overlaid on helical motifs; comparison of LLL-**30aaa** with LDL-**30aaa** illustrates this. Thus, even though the LLL-isomer is disposed to α-helical conformations (blue bar) and can match sheet-turn-sheet motifs almost as well (red bar in Figure 6.2a), the LDL-form is bias toward extended motifs and *not* helical ones (Figure 6.2b). The second observation is that correlations of mimic stereochemistries and conformational biases are beyond what the human mind can perceive; systematic data mining (the EKOS strategy) is essential for this.

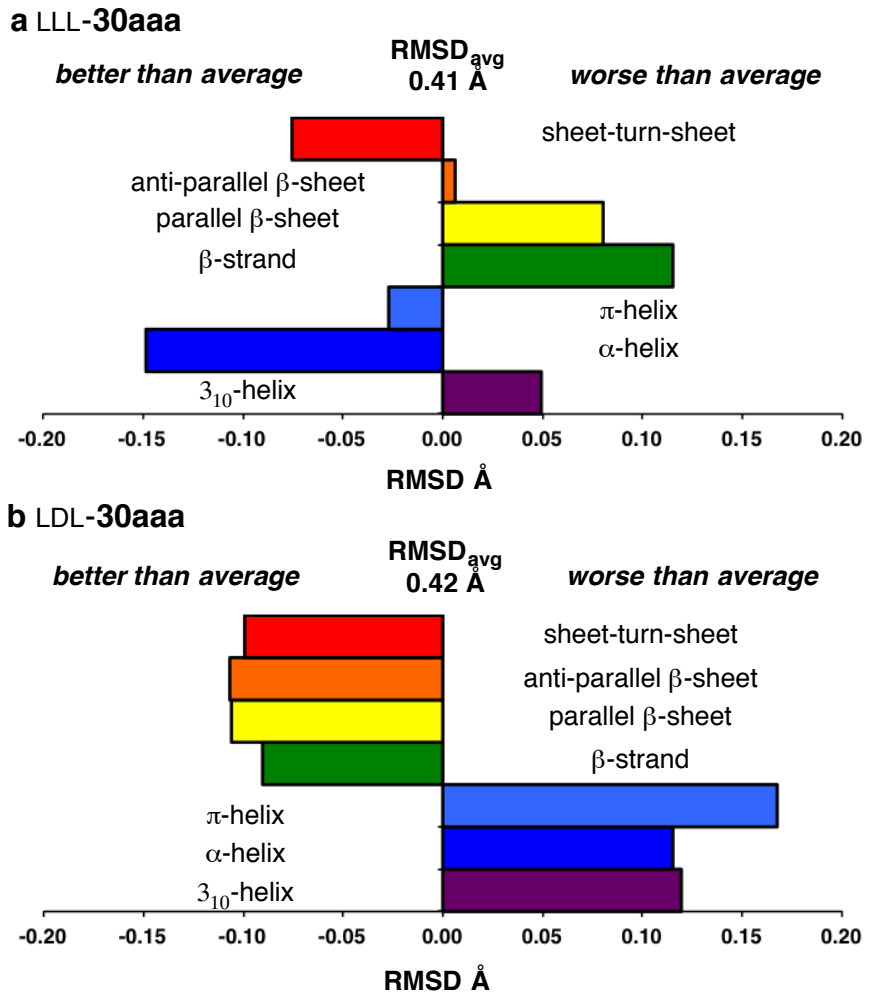


Figure 6.2 RMSD (Å) for the simulated conformers in the ensemble that best overlay the indicated ideal secondary structures, relative to the average values for the best conformers overlaid on each of the seven motifs are shown.

Most of the compounds **30** prepared here were not solids, though in one case LLL-**30fii** we were able to collect crystals and obtain an X-ray structure. That molecule crystallized in *two* similar conformations (differing by RMSD 0.28 Å, based on the 6 *C*α and *C*β coordinates, see Appendix F); but which nevertheless project the side-chains in slightly different orientations. We recently outlined another technique based on exploring key orientations that can be used to relate X-ray structures to simulated solution conformations: EKOX (see Chapter I).[55] Application of EKOX to LLL-**30fii** in the

crystal reveals that one conformer fits well on a one strand of a sheet-turn-sheet motif (Figure 6.3a), consistent with the predictions in Figure 6.2a (red bars, respectively); this is interesting because minimalist strand mimics are rare. However, neither of the conformers in the crystal structure overlaid particularly well on an α-helix; the best one is shown in Figure 6.3b.

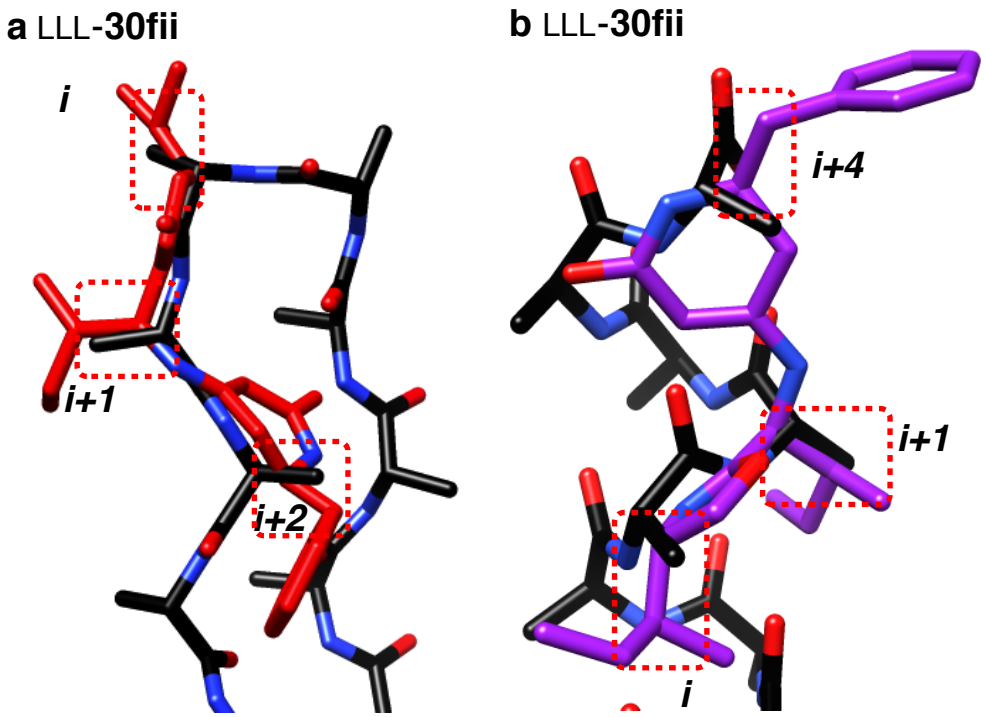


Figure 6.3 One conformer in the crystal structure of LLL-**30fii** optimally overlays with a sheet-turn-sheet motif (**a**) while neither overlays particularly well on an ideal α-helix (**b** shows the best match).

6.4 Evaluation of Simulated Conformers By EKO and DSSP

When accessible conformers of LLL-**30aaa** were simulated and data mined on over 125,000 PPIs in the PDB, 257 "hits" were found. The applied definition of a hit was that

the three methyl side-chain coordinates of the mimic overlaid with three interface side-chains with an RMSD of 0.30 Å or better. The best overlay occurred on an interface sheet-turn-sheet motif as shown in Figure 6.4 (RMSD 0.12 Å).

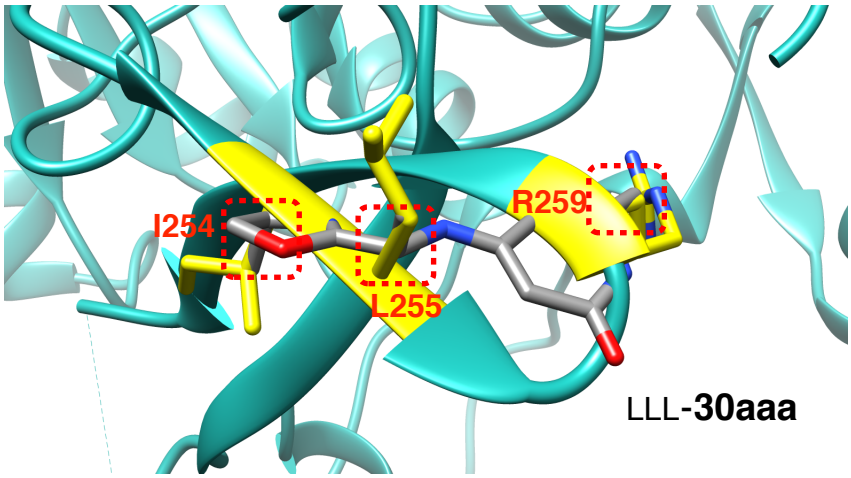


Figure 6.4 An accessible simulated conformer of LLL-**30aaa** overlays with excellent correspondence on the interface region of inosine 5'-monophosphate dehydrogenase (pdbid = 4ff0, RMSD = 0.12 Å).

Despite Figure 6.2a, it would be incorrect to assume that simulated accessible conformers of LLL-**30aaa** overlay well on only sheet-turn-sheet and helical motifs at interfaces. The statistical distribution of overlays in each of the featured ideal secondary structures is depicted in Figure 6.5. Consistent with the findings from EKOS based on *ideal* secondary structures (Figure 6.2a), α-helical and sheet-turn-sheet motifs are the most common matches for LLL-**30aaa** at interfaces. However, similar to the results observed in Chapter V, Figure 6.5 shows over 50 % of the overlays occur on nearly consecutive amino acids ("single segments" in Figure 6.5) that are not part of any ideal secondary structure. Moreover, the next most common type of overlay was on amino acid sequences from different parts of the chain, and which also do not resemble any

ideal secondary structure (*ie* "multiple segments"). Those two accounted for 87% of the "segments"; the same category we used in Chapter V. Consequently, *most of the potential for LLL-**30aaa** in interface mimicry appears not to be correlated with any particular secondary structure*.

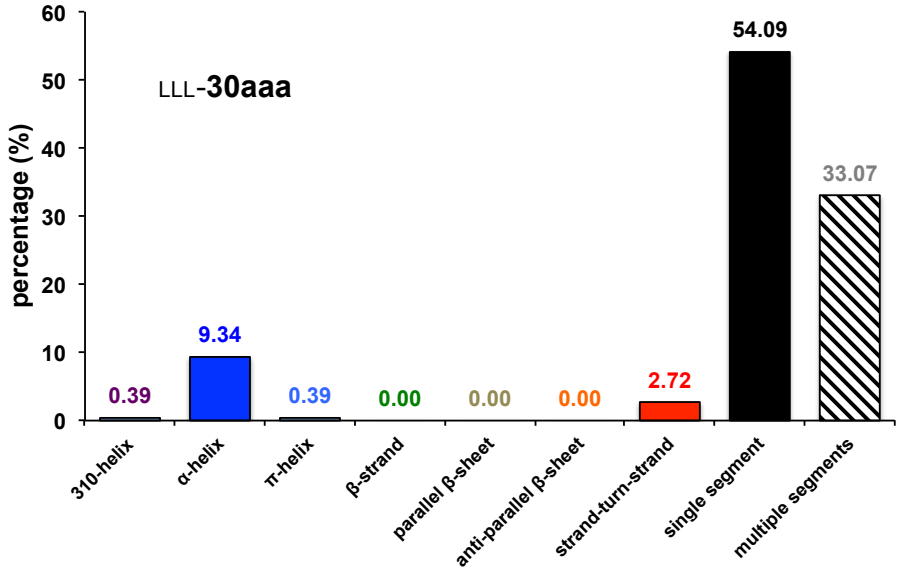


Figure 6.5 Statistical distribution of the best 257 overlays of preferred conformers of LLL-**30aaa** (generated using EKO; all RMSD <0.31 Å).

6.5 Conclusion

In this work, we introduced a novel chemotype that has the potential to be a good secondary structure mimic. By incorporating oxopiperazine moiety into the framework, this molecule showed a good helical mimic especially LLL-isomer. This molecule can be synthesized by condensing two fragments, oxazoline and oxopiperazine. Thus, the synthetic route is divergent−convergent, in which any pairs of oxazoline and oxopiperazine can be condensed to provide a variety of mimic **30**. Some parameters of

scaffolds **30** relevant to their use as cellular probes were also considered. Molecule **30aaa** has a low molecular mass and no amide bonds; these characteristics are favorable for cellular- and oral-permeability. QikProp[193,194] was used to predict some other key parameters of **30aaa**, (Figure 6.1). Simulated permeability of **30aaa** through Caco cells is excellent (*ie*, > 500 nm/s), and the estimated log octanol/water partition coefficient, 1.54, is near the mid-point of the optimal range (-2.0 to 6.5). Moreover, there are no rule of five[195] violations for this structure. Obviously, these properties will be modulated when side-chains other than methyl are involved, but the scaffold provides a good framework for probe development.

Overall, we conclude that compounds based on LLL-**30aaa** can be excellent helical mimics, but they may adopt a range of conformations that overlay well on other secondary structures, notably sheet-turn-sheet motifs. Like many other minimalist mimics, however, molecules **30** can overlay on diverse interface regions, most of which are not directly related to secondary structures.

CHAPTER VII

CONCLUSION

Previous projects led by a former graduate student in Dr. Burgess group, Dr. KO, were to develop the Exploring Key Orientations (EKO) to find best matches between three $C\alpha$-$C\beta$ side-chain vectors of semi-rigid small molecules and the side-chain vectors at PPI interfaces. This approach was validated experimentally by discovering novel small molecules disrupting HIV-1 dimerization, and later by small molecules perturbing antithrombin oligomerization conducted by Dr. Xin. However, only these two examples are not adequate to claim EKO can be used in general. Consequently, my projects were not only to synthesize novel small molecules to disrupt medicinally relevant PPIs, but also to help validate EKO as a tool to identify PPI inhibitors.

My initial project was to synthesize hydantoin-oxazoline scaffold; this structure is fascinating insofar as its conformers overlay well on many types of ideal secondary structures, and on PPI interfaces that may have distorted structure motifs. Specifically, we coined this kind of molecule as “universal peptidomimetics”. Moreover, this chemotype has a contiguous side-chains characteristic and a few numbers of degrees of freedom, which are rarely found in other peptidomimetics disrupting PPIs. Apart from hydantoin-oxazoline scaffold, other similar derivatives, triazole-oxazole and triazole-oxazoline, were developed, and those three chemotypes were tested their inhibitions against Nef•MHC-I•AP1 and NEDD8•NAE interactions (Chapter II). Although, the outcomes did not reach our expectations, we learned a couple criteria to help increase success rate of disrupting PPI interactions; (i) chemotypes should be easily accessible in order to accumulate enough compounds within a reasonable time for biological tests, (ii)

EKO does not determine binding energies or steric clashes between proteins and ligands, so other calculation methods are necessary to enhance the quality of predictions, (iii) target selection is critical because large PPI interface areas and tight complexes are difficult to inhibit interactions.

Based on experiences from Nef•MHC-I•AP1 and NEDD8•NAE projects, we moved to another target that was fit into our criteria as described above. First, we chose a new PPI target, PCSK9•LDLR, which has relatively small PPI interface area and weak binding affinity. Additionally, disrupting this complex has been validated as a target for medicinal chemistry by two FDA approved antibody drugs with no serious side effects. Thus, this is a suitable target for peptidomimetics disrupting PPI. Meanwhile, we developed a new chemotype, hydantoin-piperazine, that can be prepared via solid-phase synthesis, i.e. this chemotype is accessible by parallel syntheses to obtain a small library within short time. A key feature within this chemotype is a new substructure "pip-acids", in which we introduced piperazine scaffolds onto amino acids. The benefits behind these pip-acids are to increase rigidity of molecule relative to oligopeptides, and to add a piperazine motif found in many pharmaceutical drugs. However, the first-generation hydantoin-piperazine inhibitors, which compromised criteria (i) and (iii), did not show significant inhibitions in the first trial, so we used a docking program "Glide" to refine our structures (criteria ii). By virtually adding amino acids at *C*-terminus, we obtained a second-generation hydantoin-piperazine library, in which showed improvements of docking scores comparing to the first generation. Results from biological assays showed half of these second-generation inhibitors could disrupt PCSK9•LDLR interaction. The lead compound, LDLL-**16dlnr**, significantly increased cLDL uptake in the hepatocytes, and had measurable binding affinity toward PCSK9. This is a proof of concept that we successfully improve inhibitory activities by combining

EKO and Glide routines, and this strategy had a hit rate (~50%) higher than conventional high-throughput screening.

PCSK9•LDLR project is now starting the second phase to further improve efficiencies and bioavailability of lead compounds by a postdoctoral researcher in Dr. Burgess group. Our plan is to utilize Glide and structure-activity relationship (SAR) to make compounds less peptidic while preserving key functional groups. Consequently, I expect new scaffolds that can surpass LDLL-**16dInr** in term of better cLDL uptake and ADME properties. Besides, pip-acids are valuable as a novel building block for synthesizing other small molecules. There are still opportunities to use them for many applications; this is a main project of another graduate student in Dr. Burgess group.

At this stage, we intensively used EKOS to roughly determine how good chemotypes adopt their conformations to resemble ideal secondary structures, and EKO to search for the best overlays at PPI interfaces. However, the relationship of results between EKOS and EKO has never been explored systematically. Our study in Chapter V suggested that a good secondary structure mimicry tends to overlay more frequently on PPI interfaces within a given RMSD cutoff than a chemotype that does not match well on ideal secondary structure motifs. Even though a good secondary structure mimicry had several hits on PPI interfaces, they mostly matched on non-ideal structures. Thus, secondary structure mimicry is only an indirect indicator for interface mimicry. Moreover, "chain periodicity" of chemotype is important for designing good secondary structure mimicry. Excellent secondary structure matching usually comes from a chemotype with $C\alpha$-$C\alpha$ distance close to the oligopeptide found in ideal secondary structures or at protein interfaces. Other criteria such as rigid-core structure, minimum number of degrees of freedom, and bearing three side-chain residues are important to minimize

entropy loss and increase specificity when chemotypes dock on protein targets. Again, the results from EKO should be further refined by docking programs, e.g. Glide.

Another separate project in Chapter VI, I incorporated with Dr. Li, a visiting postdoctoral researcher in Dr. Burgess group, to prepare and analyze oxazoline-piperidine-2,4-dione minimalist mimics. These compounds could be synthesized from two separate fragments, and then condensed them together; i.e. the synthetic route is divergent-convergent. Although this synthesis strategy may not be comparable with solid-phase syntheses in term of time expenditure, condensing two set of separate fragments yields a variety of different products. Evaluation of preferred conformers showed this mimic, especially LLL-isomer, matched well on helical motif. Unfortunately, this compound has not been tested with any protein targets. I expect this new minimalist mimic can inhibit some interesting PPI targets.

REFERENCES

(1) Scott, D. E.; Bayly, A. R.; Abell, C.; Skidmore, J. *Nat. Rev. Drug Discov.* **2016**, *15*, 533-550.

(2) London, N.; Raveh, B.; Schueler-Furman, O. *Curr. Opin. Chem. Biol.* **2013**, *17*, 952-959.

(3) Surade, S.; Blundell, T. L. *Chem. Biol.* **2012**, *19*, 42-50.

(4) Bogan, A. A.; Thorn, K. S. *J. Mol. Biol.* **1998**, *280*, 1-9.

(5) Keskin, O.; Gursoy, A.; Ma, B.; Nussinov, R. *Chem. Rev.* **2008**, *108*, 1225-1244.

(6) Conte, L. L.; Chothia, C.; Janin, J. *J. Mol. Biol.* **1999**, *285*, 2177-2198.

(7) Wells, J. A.; McClendon, C. L. *Nature* **2007**, *450*, 1001-1009.

(8) Hajduk, P. J.; Greer, J. *Nat. Rev. Drug Discovery* **2007**, *6*, 211-219.

(9) Bienstock, R. J. *Curr. Pharm. Des.* **2012**, *18*, 1240-1254.

(10) Perez de Vega, M. J.; Martin-Martinez, M.; Genzalez-Muniz, R. *Curr. Top. Med. Chem.* **2007**, *7*, 33-62.

(11) Pelay-Gimeno, M.; Glas, A.; Koch, O.; Grossmann, T. N. *Angew. Chem. Int. Ed.* **2015**, *54*, 8896-8927.

(12) Milroy, L.-G.; Grossmann, T. N.; Hennig, S.; Brunsveld, L.; Ottmann, C. *Chem. Rev.* **2014**, *114*, 4695-4748.

(13) Clackson, T.; Wells, J. A. *Science* **1995**, *267*, 383-386.

(14) Yu, C.; Taylor, J. W. *Bioorg. Med. Chem.* **1999**, *7*, 161-175.

(15) Yu, C.; Taylor, J. W. *Tetrahedron Lett.* **1996**, *37*, 1731-1734.

(16) Winter, A.; Higueruelo, A. P.; Marsh, M.; Sigurdardottir, A.; Pitt, W. R.; Blundell, T. L. *Q. Rev. Biophys.* **2012**, *45*, 383-426.

(17) Osapay, G.; Taylor, J. W. *J. Am. Chem. Soc.* **1992**, *114*, 6966-6973.

(18) Osapay, G.; Taylor, J. W. *J. Am. Chem. Soc.* **1990**, *112*, 6046-6051.

(19) Higueruelo, A. P.; Jubb, H.; Blundell, T. L. *Curr. Opin. Pharmacol.* **2013**, *13*, 1-6.

(20) Walensky, L. D.; Bird, G. H. *J. Med. Chem.* **2014**, *57*, 6275-6288.

(21) Cromm, P. M.; Spiegel, J.; Grossmann, T. N. *ACS Chem. Biol.* **2015**, *10*, 1362-1375.

(22) Tan, Y. S.; Lane, D. P.; Verma, C. S. *Drug Discov. Today* **2016**, *21*, 1642-1653.

(23) Halab, L.; Gosselin, F.; Lubell, W. D. *Biopolymers* **2000**, *55*, 101-122.

(24) Wipf, P.; Xiao, J.; Stephenson, C. R. J. *Chimia* **2009**, *63*, 764-775.

(25) Ross, N. T.; Katt, W. P.; Hamilton, A. D. *Philos. Trans. R. Soc. A* **2010**, *368*, 989-1008.

(26) Watkins, A. M.; Arora, P. S. *Eur. J. Med. Chem.* **2015**, *94*, 480-488.

(27) Orner, B. P.; Ernst, J. T.; Hamilton, A. D. *J. Am. Chem. Soc.* **2001**, *123*, 5382-5383.

(28) Kutzki, O.; Park, H. S.; Ernst, J. T.; Orner, B. P.; Yin, H.; Hamilton, A. D. *J. Am. Chem. Soc.* **2002**, *124*, 11838-11839.

(29) Yin, H.; Lee, G.; Sedey, K. A.; Kutzki, O.; Park, H. S.; Orner, B. P.; Ernst, J. T.; Wang, H.-G.; Sebti, S. M.; Hamilton, A. D. *J. Am. Chem. Soc.* **2005**, *127*, 10191-10196.

(30) Kim, I. C.; Hamilton, A. D. *Org. Lett.* **2006**, *8*, 1751-1754.

(31) Ahn, J.-M.; Han, S.-Y. *Tetrahedron Lett.* **2007**, *48*, 3543-3547.

(32) Shaginian, A.; Whitby, L. R.; Hong, S.; Hwang, I.; Farooqi, B.; Searcey, M.; Chen, J.; Vogt, P. K.; Boger, D. L. *J. Am. Chem. Soc.* **2009**, *131*, 5564-5572.

(33) Plante, J.; Campbell, F.; Malkova, B.; Kilner, C.; Warriner, S. L.; Wilson, A. J. *Org. Biomol. Chem.* **2008**, *6*, 138-146.

(34) Volonterio, A.; Moisan, L.; Rebek, J., Jr. *Org. Lett.* **2007**, *9*, 3733-3736.

(35) Moisan, L.; Odermatt, S.; Gombosuren, N.; Carella, A.; Rebek, J., Jr. *Eur. J. Org. Chem.* **2008**, 1673-1676.

(36) Maity, P.; Koenig, B. *Org. Lett.* **2008**, *10*, 1473-1476.

(37) Marimganti, S.; Cheemala, M. N.; Ahn, J.-M. *Org. Lett.* **2009**, *11*, 4418-4421.

(38) Tosovska, P.; Arora, P. S. *Org. Lett.* **2010**, *12*, 1588-1591.

(39) Lee Ji, H.; Zhang, Q.; Jo, S.; Chai Sergio, C.; Oh, M.; Im, W.; Lu, H.; Lim, H.-S. *J. Am. Chem. Soc.* **2011**, *133*, 676-679.

(40) Smith, A. B., III; Charnley, A. K.; Hirschmann, R. *Acc. Chem. Res.* **2011**, *44*, 180-193.

(41) Yin, H.; Lee, G.-i.; Park, H. S.; Payne, G. A.; Rodriguez, J. M.; Sebti, S. M.; Hamilton, A. D. *Angew. Chem. Int. Ed.* **2005**, *44*, 2704-2707.

(42) Fuller, J. C.; Jackson, R. M.; Edwards, T. A.; Wilson, A. J.; Shirts, M. R. *PLoS One* **2012**, *7*, e43253.

(43) Yap, J. L.; Cao, X.; Vanommeslaeghe, K.; Jung, K.-Y.; Peddaboina, C.; Wilder, P. T.; Nan, A.; MacKerell, A. D.; Smythe, W. R.; Fletcher, S. *Org. Biomol. Chem.* **2012**, *10*, 2928-2933.

(44) Ravindranathan, P.; Lee, T.-K.; Yang, L.; Centenera, M. M.; Butler, L.; Tilley, W. D.; Hsieh, J.-T.; Ahn, J.-M.; Raj, G. V. *Nat. Commun.* **2013**, *4*, 1-11.

(45) Lao, B. B.; Drew, K.; Guarracino, D. A.; Brewer, T. F.; Heindel, D. W.; Bonneau, R.; Arora, P. S. *J. Am. Chem. Soc.* **2014**, *136*, 7877-7888.

(46) Lao, B. B.; Grishagin, I.; Mesallati, H.; Brewer, T. F.; Olenyuk, B. Z.; Arora, P. S. *Proc. Natl. Acad. Sci.* **2014**, *111*, 7531-7536.

(47) Farmer, P. S.; Ariens, E. J. *Trends Pharmacol. Sci.* **1982**, *3*, 362-365.

(48) Peczuh, M. W.; Hamilton, A. D. *Chem. Rev.* **2000**, *100*, 2479-2494.

(49) Eguchi, M.; Kahn, M. *Mini-Rev. Med. Chem.* **2002**, *2*, 447-462.

(50) Jayatunga, M. K. P.; Thompson, S.; Hamilton, A. D. *Bioorg. Med. Chem. Lett.* **2014**, *24*, 717-724.

(51) Modell, A. E.; Blosser, S. L.; Arora, P. S. *Trends Pharmacol. Sci.* **2016**, *37*, 702-713.

(52) Ko, E.; Liu, J.; Burgess, K. *Chem. Soc. Rev.* **2011**, *40*, 4411-4421.

(53) Ko, E.; Liu, J.; Perez, L. M.; Lu, G.; Schaefer, A.; Burgess, K. *J. Am. Chem. Soc.* **2011**, *133*, 462-477.

(54) Xin, D.; Ko, E.; Perez, L. M.; Ioerger, T. R.; Burgess, K. *Org. Biomol. Chem.* **2013**, *11*, 7789-7801.

(55) Xin, D.; Perez, L. M.; Ioerger, T. R.; Burgess, K. *Angew. Chem. Int. Ed.* **2014**, *53*, 3594-3598.

(56) Pettitt, B. M.; Matsunaga, T.; Al-Obeidi, F.; Gehrig, C.; Hruby, V. J.; Karplus, M. *Biophys. J.* **1991**, *60*, 1540-1544.

(57) O'Connor, S. D.; Smith, P. E.; Al-Obeidi, F.; Pettitt, B. M. *J. Med. Chem.* **1992**, *35*, 2870-2881.

(58) Ko, E.; Raghuraman, A.; Perez, L. M.; Ioerger, T. R.; Burgess, K. *J. Am. Chem. Soc.* **2013**, *135*, 167-173.

(59) Xin, D.; Holzenburg, A.; Burgess, K. *Chem. Sci.* **2014**, *5*, 4914-4921.

(60) Wipf, P. *Chemical Reviews* **1995**, *95*, 2115-2134.

(61) Tilvi, S.; Singh, K. S. *Current Organic Chemistry* **2016**.

(62) Vagner, J.; Qu, H.; Hruby, V. J. *Curr. Opin. Chem. Biol.* **2008**, *12*, 292-296.

(63) Gunning, P. T.; Glenn, M. P.; Siddiquee, K. A. Z.; Katt, W. P.; Masson, E.; Sebti, S. M.; Turkson, J.; Hamilton, A. D. *ChemBioChem* **2008**, *9*, 2800-2803.

(64) Siddiquee, K. A. Z.; Gunning, P. T.; Glenn, M.; Katt, W. P.; Zhang, S.; Schroeck, C.; Sebti, S. M.; Jove, R.; Hamilton, A. D.; Turkson, J. *ACS Chemical Biology* **2007**, *2*, 787-798.

(65) Andersson, I. E.; Batsalova, T.; Dzhambazov, B.; Edvinsson, L.; Holmdahl, R.; Kihlberg, J.; Linusson, A. *Org. Biomol. Chem.* **2010**, *8*, 2931-2940.

(66) Engelman, A.; Cherepanov, P. *Nat. Rev. Microbiol.* **2012**, *10*, 279-290.

(67) Stevenson, M. *Nat. Med.* **2003**, *9*, 853-860.

(68) Peterlin, B. M.; Trono, D. *Nat. Rev. Immunol.* **2003**, *3*, 97-107.

(69) Frankel, A. D.; Young, J. A. *Annu. Rev. Biochem.* **1998**, *67*, 1-25.

(70) Pommier, Y.; Johnson, A. A.; Marchand, C. *Nat. Rev. Drug Discov.* **2005**, *4*, 236-248.

(71) Cihlar, T.; Ray, A. S. *Antiviral Res.* **2010**, *85*, 39-58.

(72) Ghosh, A. K.; Osswald, H. L.; Prato, G. *J. Med. Chem.* **2016**, *59*, 5172-5208.

(73) Gu, S. X.; Xue, P.; Ju, X. L.; Zhu, Y. Y. *Bioorg. Med. Chem.* **2016**, *24*, 5007-5016.

(74) Blanco, J. L.; Whitlock, G.; Milinkovic, A.; Moyle, G. *Expert Opin. Pharmacother* **2015**, *16*, 1313-1324.

(75) Smithgall, T. E.; Thomas, G. *Drug Discov. Today Technol.* **2013**, *10*, e523-529.

(76) Betzi, S.; Restouin, A.; Opi, S.; Arold, S. T.; Parrot, I.; Guerlesquin, F.; Morelli, X.; Collette, Y. *Proc. Natl. Acad. Sci.* **2007**, *104*, 19256-19261.

(77) Dikeakos, J. D.; Atkins, K. M.; Thomas, L.; Emert-Sedlak, L.; Byeon, I. J.; Jung, J.; Ahn, J.; Wortman, M. D.; Kukull, B.; Saito, M.; Koizumi, H.; Williamson, D. M.; Hiyoshi, M.; Barklis, E.; Takiguchi, M.; Suzu, S.; Gronenborn, A. M.; Smithgall, T. E.; Thomas, G. *Mol. Biol. Cell* **2010**, *21*, 3279-3292.

(78) Oneyama, C.; Nakano, H.; Sharma, S. V. *Oncogene* **2002**, *21*, 2037-2050.

(79) Emert-Sedlak, L. A.; Narute, P.; Shu, S. T.; Poe, J. A.; Shi, H.; Yanamala, N.; Alvarado, J. J.; Lazo, J. S.; Yeh, J. I.; Johnston, P. A.; Smithgall, T. E. *Chem. Biol.* **2013**, *20*, 82-91.

(80) Iyer, P. C.; Zhao, J.; Emert-Sedlak, L. A.; Moore, K. K.; Smithgall, T. E.; Day, B. W. *Bioorg. Med. Chem. Lett.* **2014**, *24*, 1702-1706.

(81) Emert-Sedlak, L. A.; Loughran, H. M.; Shi, H.; Kulp, J. L., 3rd; Shu, S. T.; Zhao, J.; Day, B. W.; Wrobel, J. E.; Reitz, A. B.; Smithgall, T. E. *Bioorg. Med. Chem. Lett.* **2016**, *26*, 1480-1484.

(82) Collins, K. L.; Chen, B. K.; Kalams, S. A.; Walker, B. D.; Baltimore, D. *Nature* **1998**, *391*, 397-401.

(83) Kirchhoff, F.; Schindler, M.; Specht, A.; Arhel, N.; Muench, J. *Cell. Mol. Life Sci.* **2008**, *65*, 2621-2636.

(84) Jia, X.; Singh, R.; Homann, S.; Yang, H.; Guatelli, J.; Xiong, Y. *Nat. Struct. Mol. Biol.* **2012**, *19*, 701-706.

(85) Singh Rajendra, K.; Lau, D.; Noviello Colleen, M.; Ghosh, P.; Guatelli John, C. *PLoS One* **2009**, *4*, e8364.

(86) Soucy, T. A.; Dick, L. R.; Smith, P. G.; Milhollen, M. A.; Brownell, J. E. *Genes Cancer* **2010**, *1*, 708-716.

(87) Soucy, T. A.; Smith, P. G.; Rolfe, M. *Clin. Cancer Res.* **2009**, *15*, 3912-3916.

(88) Kumar, S.; Yoshida, Y.; Noda, M. *Biochemical and Biophysical Research Communications* **1993**, *195*, 393-399.

(89) Kamitani, T.; Kito, K.; Nguyen, H. P.; Yeh, E. T. *J. Biol. Chem.* **1997**, *272*, 28557-28562.

(90) Milhollen, M. A.; Narayanan, U.; Soucy, T. A.; Veiby, P. O.; Smith, P. G.; Amidon, B. *Cancer Res.* **2011**, *71*, 3042-3051.

(91) Nawrocki Steffan, T.; Griffin, P.; Kelly Kevin, R.; Carew Jennifer, S. *Expert Opin. Investig. Drugs* **2012**, *21*, 1563-1573.

(92) Watson, I. R.; Irwin, M. S.; Ohh, M. *Cancer Cell* **2011**, *19*, 168-176.

(93) Walden, H.; Podgorski, M. S.; Huang, D. T.; Miller, D. W.; Howard, R. J.; Minor, D. L., Jr.; Holton, J. M.; Schulman, B. A. *Mol. Cell* **2003**, *12*, 1427-1437.

(94) Walden, H.; Podgorski, M. S.; Schulman, B. A. *Nature* **2003**, *422*, 330-334.

(95) Huang, D. T.; Hunt, H. W.; Zhuang, M.; Ohi, M. D.; Holton, J. M.; Schulman, B. A. *Nature* **2007**, *445*, 394-398.

(96) Enchev, R. I.; Schulman, B. A.; Peter, M. *Nat. Rev. Mol. Cell Biol.* **2015**, *16*, 30-44.

(97) Souphron, J.; Waddell, M. B.; Paydar, A.; Tokgoz-Gromley, Z.; Roussel, M. F.; Schulman, B. A. *Biochemistry* **2008**, *47*, 8961-8969.

(98) Soucy, T. A.; Smith, P. G.; Milhollen, M. A.; Berger, A. J.; Gavin, J. M.; Adhikari, S.; Brownell, J. E.; Burke, K. E.; Cardin, D. P.; Critchley, S.; Cullis, C. A.; Doucette, A.; Garnsey, J. J.; Gaulin, J. L.; Gershman, R. E.; Lublinsky, A. R.; McDonald, A.; Mizutani, H.; Narayanan, U.; Olhava, E. J.; Peluso, S.; Rezaei, M.; Sintchak, M. D.; Talreja, T.; Thomas, M. P.; Traore, T.; Vyskocil, S.; Weatherhead, G. S.; Yu, J.; Zhang, J.; Dick, L. R.; Claiborne, C. F.; Rolfe, M.; Bolen, J. B.; Langston, S. P. *Nature* **2009**, *458*, 732-736.

(99) Swords, R. T.; Erba, H. P.; DeAngelo, D. J.; Bixby, D. L.; Altman, J. K.; Maris, M.; Hua, Z.; Blakemore, S. J.; Faessel, H.; Sedarati, F.; Dezube, B. J.; Giles, F. J.; Medeiros, B. C. *Br. J. Haematol.* **2015**, *169*, 534-543.

(100) Lu, P.; Liu, X.; Yuan, X.; He, M.; Wang, Y.; Zhang, Q.; Ouyang, P.-k. *ACS Chem. Biol.* **2016**, *11*, 1901-1907.

(101) Ma, H.; Zhuang, C.; Xu, X.; Li, J.; Wang, J.; Min, X.; Zhang, W.; Zhang, H.; Miao, Z. *Eur. J. Med. Chem.* **2017**, *133*, 174-183.

(102) Fukuyama, T.; Tokuyama, H. *Aldrichimica Acta* **2004**, *37*, 87-96.

(103) Morwick, T.; Hrapchak, M.; DeTuri, M.; Campbell, S. *Org. Lett.* **2002**, *4*, 2665-2668.

(104) Alper, P. B.; Hung, S.-C.; Wong, C.-H. *Tetrahedron Lett.* **1996**, *37*, 6029-6032.

(105) Kolb, H. C.; Finn, M. G.; Sharpless, K. B. *Angew. Chem. Int. Ed.* **2001**, *40*, 2004-2021.

(106) Eglen, R. M.; Reisine, T.; Roby, P.; Rouleau, N.; Illy, C.; Bosse, R.; Bielefeld, M. *Curr. Chem. Genomics* **2008**, *1*, 2-10.

(107) Meusel, M.; Guetschow, M. *Org. Prep. Proced. Int.* **2004**, *36*, 391-443.

(108) Yamaguchi, J.-I.; Harada, M.; Kondo, T.; Noda, T.; Suyama, T. *Chem. Lett.* **2003**, *32*, 372-373.

(109) Zhang, D.; Xing, X.; Cuny, G. D. *J. Org. Chem.* **2006**, *71*, 1750-1753.

(110) Miller, J. J.; Rajaram, S.; Pfaffenroth, C.; Sigman, M. S. *Tetrahedron* **2009**, *65*, 3110-3119.

(111) Swords, R. T.; Kelly, K. R.; Smith, P. G.; Garnsey, J. J.; Mahalingam, D.; Medina, E.; Oberheu, K.; Padmanabhan, S.; O'Dwyer, M.; Nawrocki, S. T.; Giles, F. J.; Carew, J. S. *Blood* **2010**, *115*, 3796-3800.

(112) Nawrocki, S. T.; Kelly, K. R.; Smith, P. G.; Espitia, C. M.; Possemato, A.; Beausoleil, S. A.; Milhollen, M.; Blakemore, S.; Thomas, M.; Berger, A.; Carew, J. S. *Clin. Cancer Res.* **2013**, *19*, 3577-3590.

(113) Costet, P.; Krempf, M.; Cariou, B. *Trends Biochem. Sci.* **2008**, *33*, 426-434.

(114) Li, J.; Tumanut, C.; Gavigan, J.-A.; Huang, W.-J.; Hampton, E. N.; Tumanut, R.; Suen, K. F.; Trauger, J. W.; Spraggon, G.; Lesley, S. A.; Liau, G.; Yowe, D.; Harris, J. L. *Biochem. J.* **2007**, *406*, 203-207.

(115) McNutt, M. C.; Lagace, T. A.; Horton, J. D. *J. Biol. Chem.* **2007**, *282*, 20799-20803.

(116) Norata Giuseppe, D.; Tibolla, G.; Catapano Alberico, L. *Annu. Rev. Pharmacol. Toxicol.* **2014**, *54*, 273-293.

(117) Tveten, K.; Holla, O. L.; Cameron, J.; Strom, T. B.; Berge, K. E.; Laerdahl, J. K.; Leren, T. P. *Hum. Mol. Genet.* **2012**, *21*, 1402-1409.

(118) Yamamoto, T.; Lu, C.; Ryan, R. O. *J. Biol. Chem.* **2011**, *286*, 5464-5470.

(119) Holla, O. L.; Cameron, J.; Tveten, K.; Stroem, T. B.; Berge, K. E.; Laerdahl, J. K.; Leren, T. P. *J. Lipid Res.* **2011**, *52*, 1787-1794.

(120) Zhang, D.-W.; Lagace, T. A.; Garuti, R.; Zhao, Z.; McDonald, M.; Horton, J. D.; Cohen, J. C.; Hobbs, H. H. *J. Biol. Chem.* **2007**, *282*, 18602-18612.

(121) Horton, J. D.; Cohen, J. C.; Hobbs, H. H. *J. Lipid Res.* **2009**, S172-S177.

(122) Seidah, N. G. *Expert Opin. Ther. Targets* **2009**, *13*, 19-28.

(123) Cohen, J. C.; Boerwinkle, E.; Mosley, T. H., Jr.; Hobbs, H. H. *N. Engl. J. Med.* **2006**, *354*, 1264-1272.

(124) Abifadel, M.; Varret, M.; Rabes, J.-P.; Allard, D.; Ouguerram, K.; Devillers, M.; Cruaud, C.; Benjannet, S.; Wickham, L.; Erlich, D.; Derre, A.; Villeger, L.; Farnier, M.; Beucler, I.; Bruckert, E.; Chambaz, J.; Chanu, B.; Lecerf, J.-M.; Luc, G.; Moulin, P.; Weissenbach, J.; Prat, A.; Krempf, M.; Junien, C.; Seidah, N. G.; Boileau, C. *Nat. Genet.* **2003**, *34*, 154-156.

(125) Zhao, Z.; Tuakli-Wosornu, Y.; Lagace, T. A.; Kinch, L.; Grishin, N. V.; Horton, J. D.; Cohen, J. C.; Hobbs, H. H. *Am. J. Hum. Genet.* **2006**, *79*, 514-523.

(126) Steinberg, D.; Witztum, J. L. *Proc. Natl. Acad. Sci.* **2009**, *106*, 9546-9547.

(127) Stein, E. A.; Swergold, G. D. *Curr. Atheroscler. Rep.* **2013**, *15*, 1-14.

(128) Seidah, N. G.; Prat, A. *Nat. Rev. Drug Discovery* **2012**, *11*, 367-383.

(129) Chan, J. C. Y.; Piper, D. E.; Cao, Q.; Liu, D.; King, C.; Wang, W.; Tang, J.; Liu, Q.; Higbee, J.; Xia, Z.; Di, Y.; Shetterly, S.; Arimura, Z.; Salomonis, H.; Romanow, W. G.; Thibault, S. T.; Zhang, R.; Cao, P.; Yang, X.-P.; Yu, T.; Lu, M.; Retter, M. W.; Kwon, G.; Henne, K.; Pan, O.; Tsai, M.-M.; Fuchslocher, B.; Yang, E.; Zhou, L.; Lee, K. J.; Daris, M.; Sheng, J.; Wang, Y.; Shen, W. D.; Yeh, W.-C.; Emery, M.; Walker, N. P. C.; Shan, B.; Schwarz, M.; Jackson, S. M. *Proc. Natl. Acad. Sci.* **2009**, *106*, 9820-9825.

(130) Ni Yan, G.; Di Marco, S.; Condra Jon, H.; Peterson Laurence, B.; Wang, W.; Wang, F.; Pandit, S.; Hammond Holly, A.; Rosa, R.; Cummings Richard, T.; Wood Dana,

D.; Liu, X.; Bottomley Matthew, J.; Shen, X.; Cubbon Rose, M.; Wang, S.-p.; Johns Douglas, G.; Volpari, C.; Hamuro, L.; Chin, J.; Huang, L.; Zhao Jing, Z.; Vitelli, S.; Haytko, P.; Wisniewski, D.; Mitnaul Lyndon, J.; Sparrow Carl, P.; Hubbard, B.; Carfi, A.; Sitlani, A. *J. Lipid Res.* **2011**, *52*, 78-86.

(131) Raal, F. J.; Stein, E. A.; Dufour, R.; Turner, T.; Civeira, F.; Burgess, L.; Langslet, G.; Scott, R.; Olsson, A. G.; Sullivan, D.; Hovingh, G. K.; Cariou, B.; Gouni-Berthold, I.; Somaratne, R.; Bridges, I.; Scott, R.; Wasserman, S. M.; Gaudet, D. *Lancet* **2015**, *385*, 331-340.

(132) Raal, F. J.; Honarpour, N.; Blom, D. J.; Hovingh, G. K.; Xu, F.; Scott, R.; Wasserman, S. M.; Stein, E. A. *Lancet* **2015**, *385*, 341-350.

(133) Zhang, Y.; Eigenbrot, C.; Zhou, L.; Shia, S.; Li, W.; Quan, C.; Tom, J.; Moran, P.; Di Lello, P.; Skelton, N. J.; Kong-Beltran, M.; Peterson, A.; Kirchhofer, D. *J. Biol. Chem.* **2014**, *289*, 942-955.

(134) Mayer, G.; Poirier, S.; Seidah, N. G. *J. Biol. Chem.* **2008**, *283*, 31791-31801.

(135) Seidah, N. G.; Poirier, S.; Denis, M.; Parker, R.; Miao, B.; Mapelli, C.; Prat, A.; Wassef, H.; Davignon, J.; Hajjar, K. A.; Mayer, G. *PLoS One* **2012**, *7*, e41865.

(136) Shan, L.; Pang, L.; Zhang, R.; Murgolo, N. J.; Lan, H.; Hedrick, J. A. *Biochem. Biophys. Res. Commun.* **2008**, *375*, 69-73.

(137) Du, F.; Hui, Y.; Zhang, M.; Linton, M. F.; Fazio, S.; Fan, D. *J. Biol. Chem.* **2011**, *286*, 43054-43061.

(138) Palmer-Smith, H.; Basak, A. *Curr. Med. Chem.* **2010**, *17*, 2168-2182.

(139) Lammi, C.; Zanoni, C.; Aiello, G.; Arnoldi, A.; Grazioso, G. *Sci. Rep.* **2016**, *6*, 29931.

(140) Guay, D. C., S.; Lachance, N.; CHIASSON, J.F.; Truong, V.L.; Lacombe, P.; Skorey, K.; Seidah, N.G. **2014**, WO2014139008.

(141) F. Powell, M. In *Annual Reports in Medicinal Chemistry* 1993; Vol. 28, p 285-294.

(142) Barta, T. E. B., J. W.; Monroe, K. D.; Meuhlemann, M.M. **2016**, WO2016029037.

(143) Min, D. K.; Lee, H. S.; Lee, N.; Lee, C. J.; Song, H. J.; Yang, G. E.; Yoon, D.; Park, S. W. *Yonsei Med. J.* **2015**, *56*, 1251-1257.

(144) Park, S. W. L., H.S.; Min, D.K.; Lee, N.R.; Lee, C.J. **2016**, WO2016108572.

(145) Stucchi, M.; Grazioso, G.; Lammi, C.; Manara, S.; Zanoni, C.; Arnoldi, A.; lesma, g.; Silvani, A. *Org. Biomol. Chem.* **2016**, 9736-9740.

(146) Piper, D. E.; Jackson, S.; Liu, Q.; Romanow, W. G.; Shetterly, S.; Thibault, S. T.; Shan, B.; Walker, N. P. C. *Structure* **2007**, *15*, 545-552.

(147) Hampton, E. N.; Knuth, M. W.; Li, J.; Harris, J. L.; Lesley, S. A.; Spraggon, G. *Proc. Natl. Acad. Sci.* **2007**, *104*, 14604-14609.

(148) Cunningham, D.; Danley, D. E.; Geoghegan, K. F.; Griffor, M. C.; Hawkins, J. L.; Subashi, T. A.; Varghese, A. H.; Ammirati, M. J.; Culp, J. S.; Hoth, L. R.; Mansour, M. N.; McGrath, K. M.; Seddon, A. P.; Shenolikar, S.; Stutzman-Engwall, K. J.; Warren, L. C.; Xia, D.; Qiu, X. *Nat. Struct. Mol. Biol.* **2007**, *14*, 413-419.

(149) Lo Surdo, P.; Bottomley, M. J.; Calzetta, A.; Settembre, E. C.; Cirillo, A.; Pandit, S.; Ni, Y. G.; Hubbard, B.; Sitlani, A.; Carfi, A. *EMBO Rep.* **2011**, *12*, 1300-1305.

(150) Kwon, H. J.; Lagace, T. A.; McNutt, M. C.; Horton, J. D.; Deisenhofer, J. *Proc. Natl. Acad. Sci.* **2008**, *105*, 1820-1825.

(151) Konnert, L.; Lamaty, F.; Martinez, J.; Colacino, E. *Chem. Rev.* **2017**, *117*, 13757-13809.

(152) Ware, E. *Chem. Rev.* **1950**, *46*, 403-470.

(153) Jamieson, A. G.; Russell, D.; Hamilton, A. D. *Chem. Commun. (Camb)* **2012**, *48*, 3709-3711.

(154) Vazquez, J.; Garcia-Jareno, A.; Mondragon, L.; Rubio-Martinez, J.; Perez-Paya, E.; Albericio, F. *ChemMedChem* **2008**, *3*, 979-985.

(155) Vitaku, E.; Smith, D. T.; Njardarson, J. T. *J. Med. Chem.* **2014**, *57*, 10257-10274.

(156) Al-Ghorbani, M.; Bushra, B. A.; Zabiulla; Mamatha, S. V.; Khanum, S. A. *Research Journal of Pharmacy and Technology* **2015**, *8*, 611.

(157) Shaquiquzzaman, M.; Verma, G.; Marella, A.; Akhter, M.; Akhtar, W.; Khan, M. F.; Tasneem, S.; Alam, M. M. *Eur. J. Med. Chem.* **2015**, *102*, 487-529.

(158) Restorp, P.; Rebek, J., Jr. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 5909-5911.

(159) Moisan, L.; Odermatt, S.; Gombosuren, N.; Carella, A.; Rebek, J. *Eur. J. Org. Chem.* **2008**, *10*, 1673-1676.

(160) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graph.* **1996**, *14*, 33-38, 27-38.

(161) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781-1802.

(162) McNutt, M. C.; Kwon, H. J.; Chen, C.; Chen, J. R.; Horton, J. D.; Lagace, T. A. *J. Biol. Chem.* **2009**, *284*, 10561-10570.

(163) Schroeder Christina, I.; Swedberg Joakim, E.; Akcan, M.; Clayton Daniel, J.; Daly Norelle, L.; Cheneval, O.; Colless, B.; Walsh, P.; Sunderland, P.; Withka Jane, M.; Borzilleri Kris, A.; Griffor, M.; Ammirati, M.; Liu, S.; Rosengren, K. J.; Stock, I.; Reyes, A.; Dullea, R.; McClure Kim, F.; Tu, M.; Bhattacharya Samit, K.; Liras, S.; Price David, A.; Craik David, J. *Chem. Biol.* **2014**, *21*, 284-294.

(164) Bottomley, M. J.; Cirillo, A.; Orsatti, L.; Ruggeri, L.; Fisher, T. S.; Santoro, J. C.; Cummings, R. T.; Cubbon, R. M.; Lo Surdo, P.; Calzetta, A.; Noto, A.; Baysarowich, J.; Mattu, M.; Talamo, F.; De Francesco, R.; Sparrow, C. P.; Sitlani, A.; Carfi, A. *J. Biol. Chem.* **2009**, *284*, 1313-1323.

(165) Huang, J.; Xu, W.; Xie, H.; Li, S. *J. Org. Chem.* **2012**, *77*, 7506-7511.

(166) Brown, D. G.; Bernstein, P. R.; Griffin, A.; Wesolowski, S.; Labrecque, D.; Tremblay, M. C.; Sylvester, M.; Mauger, R.; Edwards, P. D.; Throner, S. R.; Folmer, J. J.; Cacciola, J.; Scott, C.; Lazor, L. A.; Pourashraf, M.; Santhakumar, V.; Potts, W. M.;

Sydserff, S.; Giguere, P.; Levesque, C.; Dasser, M.; Groblewski, T. *J. Med. Chem.* **2014**, *57*, 733-758.

(167) Sanphui, P.; Bolla, G.; Nangia, A. *Crystal Growth & Design* **2012**, *12*, 2023-2036.

(168) Stones, G.; Tripoli, R.; McDavid, C. L.; Roux-Duplatre, K.; Kennedy, A. R.; Sherrington, D. C.; Gibson, C. L. *Org. Biomol. Chem.* **2008**, *6*, 374-384.

(169) Gao, R.; Canney, D. J. *J. Org. Chem.* **2010**, *75*, 7451-7453.

(170) Careskey, H. E.; Davis, R. A.; Alborn, W. E.; Troutt, J. S.; Cao, G.; Konrad, R. J. *J. Lipid Res.* **2008**, *49*, 394-398.

(171) Dubuc, G.; Chamberland, A.; Wassef, H.; Davignon, J.; Seidah, N. G.; Bernier, L.; Prat, A. *Arterioscler Thromb. Vasc. Biol.* **2004**, *24*, 1454-1459.

(172) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. *J. Med. Chem.* **2004**, *47*, 1739-1749.

(173) Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L. *J. Med. Chem.* **2004**, *47*, 1750-1759.

(174) Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrin, P. C.; Mainz, D. T. *J. Med. Chem.* **2006**, *49*, 6177-6196.

(175) Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. *J. Phys. Chem. B* **2001**, *105*, 6474-6487.

(176) Pagadala, N. S.; Syed, K.; Tuszynski, J. *Biophys. Rev.* **2017**, *9*, 91-102.

(177) Kan, T.; Fukuyama, T. *Chem. Commun.* **2004**, 353-359.

(178) Zhang, Y.; Zhou, L.; Kong-Beltran, M.; Li, W.; Moran, P.; Wang, J.; Quan, C.; Tom, J.; Kolumam, G.; Elliott, J. M.; Skelton, N. J.; Peterson, A. S.; Kirchhofer, D. *J. Mol. Biol.* **2012**, *422*, 685-696.

(179) Stacy, D. M.; Le Quement, S. T.; Hansen, C. L.; Clausen, J. W.; Tolker-Nielsen, T.; Brummond, J. W.; Givskov, M.; Nielsen, T. E.; Blackwell, H. E. *Org. Biomol. Chem.* **2013**, *11*, 938-954.

(180) Touw Wouter, G.; Baakman, C.; Black, J.; Krieger, E.; Vriend, G.; te Beek Tim, A. H.; Joosten Robbie, P. *Nucleic Acids Res.* **2015**, *43*, D364-D368.

(181) Kabsch, W.; Sander, C. *Biopolymers* **1983**, *22*, 2577-2637.

(182) Frishman, D.; Argos, P. *Proteins* **1995**, *23*, 566-579.

(183) Heinig, M.; Frishman, D. *Nucleic Acids Res.* **2004**, *32*, W500-502.

(184) Lao, B. B.; Grishagin, I.; Mesallati, H.; Brewer, T. F.; Olenyuk, B. Z.; Arora, P. S. *Proc. Natl. Acad. Sci.* **2014**, *111*, 7531-7536.

(185) Loughlin, W. A.; Tyndall, J. D. A.; Glenn, M. P.; Fairlie, D. P. *Chem. Rev.* **2004**, *104*, 6085-6118.

(186) Cummings, C. G.; Hamilton, A. D. *Current Opinion in Chemical Biology* **2010**, *14*, 341-346.

(187) Davis, J. M.; Truong, A.; Hamilton, A. D. *Org. Lett.* **2005**, *7*, 5405-5408.

(188) Azzarito, V.; Long, K.; Murphy, N. S.; Wilson, A. J. *Nat. Chem.* **2013**, *5*, 161-173.

(189) Rajaram, S.; Sigman, M. S. *Org. Lett.* **2002**, *4*, 3399-3401.

(190) Xin, D., K. Burgess *Org. Lett.* **2014**, *16*, 2108-2110.

(191) Marin, J.; Didierjean, C.; Aubry, A.; Casimir, J. R.; Briand, J. P.; Guichard, G. *J. Org. Chem.* **2004**, *69*, 130-141.

(192) Guichard, G.; Chaloin, O.; Cabart, F.; Marin, J.; Zhang, H.; Mihara, H. *Org. Synth.* **2008**, *85*, 147-157.

(193) Jorgensen, W. L.; Duffy, E. M. *Bioorg. Med. Chem. Lett.* **2000**, *10*, 1155-1158.

(194) Duffy, E. M.; Jorgensen, W. L. *J. Am. Chem. Soc.* **2000**, *122*, 2878-2888.

(195) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. *Adv. Drug Delivery Rev.* **1997**, *23*, 3-25.

(196) Bharate, S. S.; Vishwakarma, R. A. *Bioorg. Med. Chem. Lett.* **2015**, *25*, 1561-1567.

(197) McKennon, M. J.; Meyers, A. I.; Drauz, K.; Schwarm, M. *J. Org. Chem.* **1993**, *58*, 3568-3571.

(198) Rodriguguez, M. L., M.; Doulut S.; Heitz A.; Martinez, J. *Tetrahedron Letters* **1991**, *32*, 923-926.

(199) Mosa, F.; Thirsk, C.; Vaultier, M.; Maw, G.; Whiting, A. *Org. Synth.* **2008**, *85*, 219-230.

(200) Menichincheri, M.; Bargiotti, A.; Berthelsen, J.; Bertrand, J. A.; Bossi, R.; Ciavolella, A.; Cirla, A.; Cristiani, C.; Croci, V.; D'Alessio, R.; Fasolini, M.; Fiorentini, F.; Forte, B.; Isacchi, A.; Martina, K.; Molinari, A.; Montagnoli, A.; Orsini, P.; Orzi, F.; Pesenti, E.; Pezzetta, D.; Pillan, A.; Poggesi, I.; Roletto, F.; Scolaro, A.; Tato, M.; Tibolla, M.; Valsasina, B.; Varasi, M.; Volpi, D.; Santocanale, C.; Vanotti, E. *J. Med. Chem.* **2009**, *52*, 293-307.

(201) Fedoseyenko, D.; Raghuraman, A.; Ko, E.; Burgess, K. *Org. Biomol. Chem.* **2012**, *10*, 921-924.

(202) Hosseini, M.; Grau Jakob, S.; Sorensen Kasper, K.; Sotofte, I.; Tanner, D.; Murray, A.; Tonder Janne, E. *Org. Biomol. Chem.* **2007**, *5*, 2207-2210.

APPENDIX A

GENERAL EXPERIMENTAL PROCEDURES

General Methods

All reactions were carried out under an inert atmosphere (nitrogen, or argon where stated) with dry solvents under anhydrous conditions. Glassware for anhydrous reactions was dried in an oven at 140 ^{o}C for minimum 6 h prior to use. Dry solvents were obtained by passing the previously degassed solvents through activated alumina columns. Reagents were purchased at a high commercial quality (typically 97 % or higher) and used without further purification, unless otherwise stated. High field NMR spectra were recorded with Bruker Avance III at 400 MHz for ^{1}H, and 100 MHz for ^{13}C and were calibrated using residual non-deuterated solvent as an internal reference ($CDCl_3$: ^{1}H NMR = 7.27, ^{13}C NMR = 77.0, MeOD: ^{1}H NMR = 3.30, ^{13}C NMR = 49.0, DMSO-d_6: ^{1}H NMR = 2.50, ^{13}C NMR = 39.5). Flash chromatography was performed using silica gel (230-600 mesh). Analytical thin layer chromatography (TLC) was carried out on Merck silica gel plates with QF 254 indicator and visualized by UV, ceric ammonium molybdate, ninhydrin, *para*-methoxybenzaldehyde and/or potassium permanganate stains. The following abbreviations were used to explain the multiplicities: s = singlet, d = doublet, t = triplet, q = quartet, quint = quintet, dd = double doublet, dt = double triplet, dq = double quartet, m = multiplet, br = broad. Electrospray ionization mass spectrometry (ESI-MS) data were collected on triple-stage quadrupole instrument in a positive mode. LC-MS analyses were collected from Agilent 1260 Infinity Quaternary LC and Agilent 6120 Quadrupole LC/MS modules using Poroshell 120 EC-C18 2.7 μM (4.6 x 50 mm) column in 5-95% MeCN/water gradient with 0.1% formic acid

over 10 minutes. Microwave irradiation for solid-phase syntheses was done using CEM MARS 5® system. All statistical analyses were carried out by Graphpad Prism version 6.0 (Graphpad Software). Student's t-test was used to determine significant differences between compounds and negative control. Results are represented as means ± SD.

General Procedure For X-Ray Structure Determination

A Leica MZ 75 microscope was used to identify a suitable colorless multi-faceted crystal with very well-defined faces with dimensions (max, intermediate, and min) 0.2 mm x 0.02 mm x 0.02 mm from a representative sample of crystals of the same habit. The crystal mounted on a nylon loop was then placed in a cold nitrogen stream maintained at 110 K.

A BRUKER D8-GADDS X-ray (three-circle) diffractometer was employed for crystal screening, unit cell determination, and data collection. The goniometer was controlled using the FRAMBO software suite[1]. The sample was optically centered with the aid of a video camera such that no translations were observed as the crystal was rotated through all positions. The detector was set at 6.0 cm from the crystal sample (MWPC Hi-Star Detector, 512x512 pixel). The X-ray radiation employed was generated from a Cu sealed X-ray tube (K_α = 1.54184 Å with a potential of 40 kV and a current of 40 mA) fitted with a graphite monochromator in the parallel mode (175 mm collimator with 0.5 mm mono-capillary optics).

The rotation exposure indicated acceptable crystal quality and the unit cell determination was undertaken. 2100 data frames were taken at widths of 0.5° with an exposure time of 10 seconds. Over 6000 reflections were centered and their positions were

determined. These reflections were used in the auto-indexing procedure to determine the unit cell. A suitable cell was found and refined by nonlinear least squares and Bravais lattice procedures. No super-cell or erroneous reflections were observed.

After careful examination of the unit cell, a standard data collection procedure was initiated. This procedure consists of collection of one hemisphere of data collected using omega scans, involving the collection 0.5° frames at fixed angles for ϕ, 2θ, and χ (2θ = -28°, χ = 54.73°, 2θ = -90°, χ = 54.73°), while varying omega. Addition data frames were collected to complete the data set and for absolute structure determination. Each frame was exposed for 5 sec. The total data collection was performed for duration of approximately 24 hours at 110K. No significant intensity fluctuations of equivalent reflections were observed.

Data Reduction, Structure Solution, and Refinement

Integrated intensity information for each reflection was obtained by reduction of the data frames with the program SAINT.[2] The integration method employed a three dimensional profiling algorithm and all data were corrected for Lorentz and polarization factors, as well as for crystal decay effects. Finally, the data was merged and scaled to produce a suitable data set. The absorption correction program SADABS[3] was employed to correct the data for absorption effects. A solution was obtained readily using SHELXTL.[4] All non-hydrogen atoms were refined with anisotropic thermal parameters. The hydrogen atoms bound to carbon were placed in idealized positions [C–H = 0.96 Å, $U_{iso}(H) = 1.2 \times U_{iso}(C)$]. The structure was refined (weighted least squares refinement on F^2) to convergence.[4] X-seed was employed for the final data presentation and structure plots.[5]

[1] FRAMBO “Program for Data Collection on Area Detectors” BRUKER AXS Inc., 5465 East Cheryl Parkway, Madison, WI 53711-5373 USA

[2] SAINT, “Program for Data Reduction from Area Detectors” ” BRUKER AXS Inc., 5465 East Cheryl Parkway, Madison, WI 53711-5373 USA

[3] SADABS, Sheldrick, G.M. “Program for Absorption Correction of Area Detector Frames”, BRUKER AXS Inc., 5465 East Cheryl Parkway, Madison, WI 53711-5373 USA

[4] SHELXTL, Sheldrick, G.M. (2008). Acta Cryst. A64, 112-122

[5] OLEX2, Dolomanov, O.V., Bourhis, L.L., Gildea, R.J., Howard, J.A.K. & Puschmann, H. (2009), J. Appl. Cryst. 42, 339-341.

APPENDIX B

EXPERIMENTAL PROCEDURES FOR CHAPTER II

Table B.1 Matching results of compound 1 on Nef•MHC-I•AP1 complex.

interface	compound configuration	sequence	RMSD	mimic on
MHC-I•AP1	LDL	W408-P407-L406	0.08	AP1
	DLD	L406-P407-W408	0.19	AP1
	LLD	L406-P407-W408	0.23	AP1
	LDD	V409-W408-P407	0.31	AP1
	LDD	W408-P407-L406	0.33	AP1
	DDD	W408-P407-L406	0.34	AP1
	LLL	L406-P407-W408	0.35	AP1
Nef•AP1	LLD	F68-P69-V70	0.22	Nef
	DLL	E64-E65-V66	0.23	Nef
	DDL	V66-E65-E64	0.26	Nef
	DDD	P72-T71-V70	0.26	Nef
	LDL	V70-P69-F68	0.29	Nef
	DLL	V70-T71-P72	0.30	Nef
	DLL	F301-K302-R303	0.31	AP1
	DDD	K302-F301-Q300	0.31	AP1
	LDD	V66-E65-E64	0.32	Nef
	DLD	F301-K302-R303	0.32	AP1
	LLD	E64-E65-V66	0.33	Nef
	DDL	R303-K302-F301	0.34	AP1
	LDD	V70-P69-F68	0.34	Nef
	LLL	V70-T71-P72	0.35	Nef

Table B.1 Continued.

MHC-I	LLL	D327-S328-A329	0.14
	DLL	D327-S328-A329	0.20
	DDD	A329-S328-D327	0.20
	LDD	S321-Y320-S319	0.20
	LLD	S321-Q322-A323	0.21
	DDL	A329-S328-D327	0.21
	LDL	S328-D327-S326	0.22
	LDD	A323-Q322-S321	0.23
	LLD	S319-Y320-S321	0.24
	DLD	S326-D327-S328	0.26
	DLD	D327-S328-A329	0.30
	DDD	S328-D327-S326	0.33
	DLL	Y320-S321-Q322	0.34
	LDL	A329-S328-D327	0.34
	LLD	D327-S328-A329	0.35
	DLD	Q322-A323-A324	0.35

Table B.2. Matching results of compound **1** and **3** on *C*-terminal NEDD8

Compound	compound configuration	sequence	RMSD
1	LDL	L73-A72-L71	0.13
	LLD	L71-A72-L73	0.22
	DLD	L71-A72-L73	0.26
3	DL	L73-A72-L71	0.33

Conditions to Cyclize Oxazolines

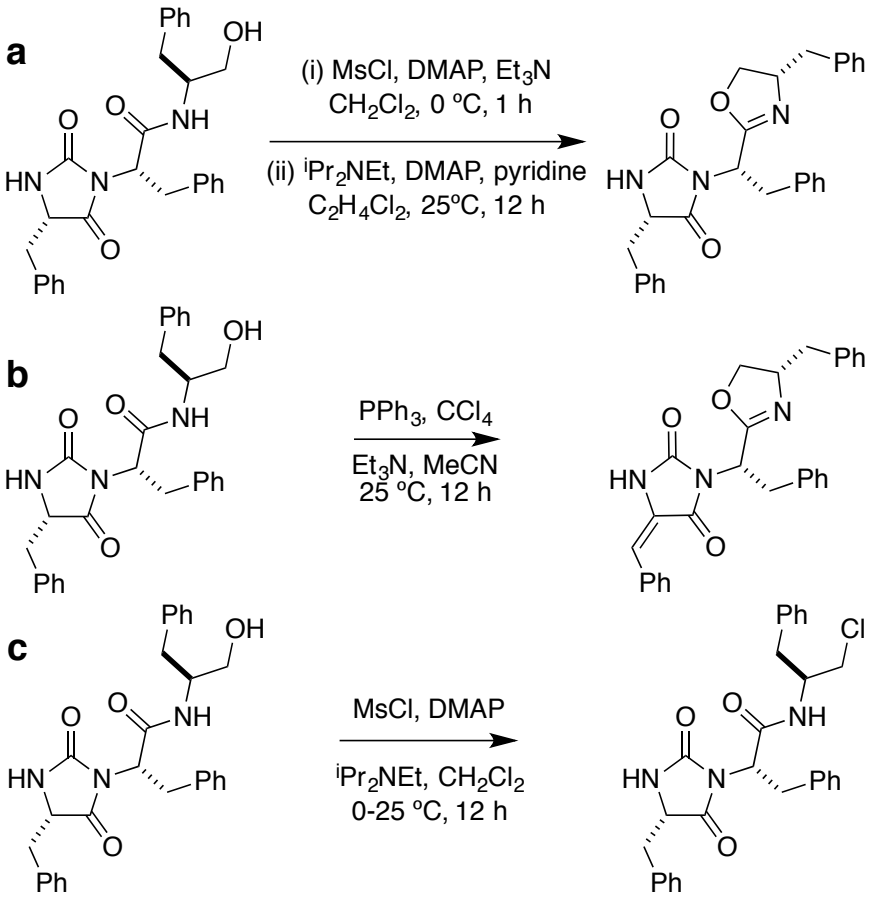
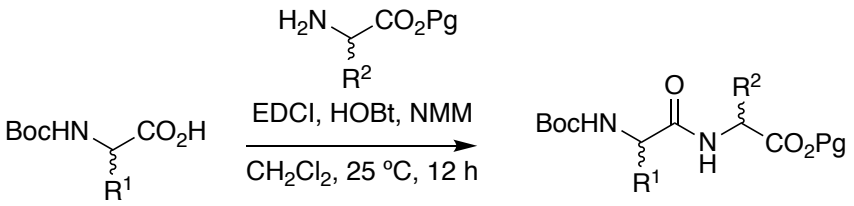


Figure B.1 **a** Successful oxazoline cyclization was performed using methanesulfonyl chloride in the 2-steps manner. **b** Condition that involved PPh_3 and CCl_4 led to the side-chain dehydrogenation at hydantoin ring. **c** Using methanesulfonyl chloride in the 1-step method led to the chlorination instead of cyclization.

General Procedure for the Preparation of Dipeptide **7**

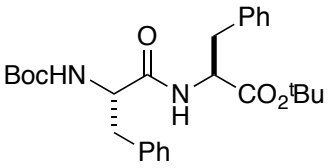


7

LL-**7ff**, 87%; DL-**7la**, 88%
LD-**7la**, 99%; LL-**7la**, 91%
LD-**7vf**, 90%

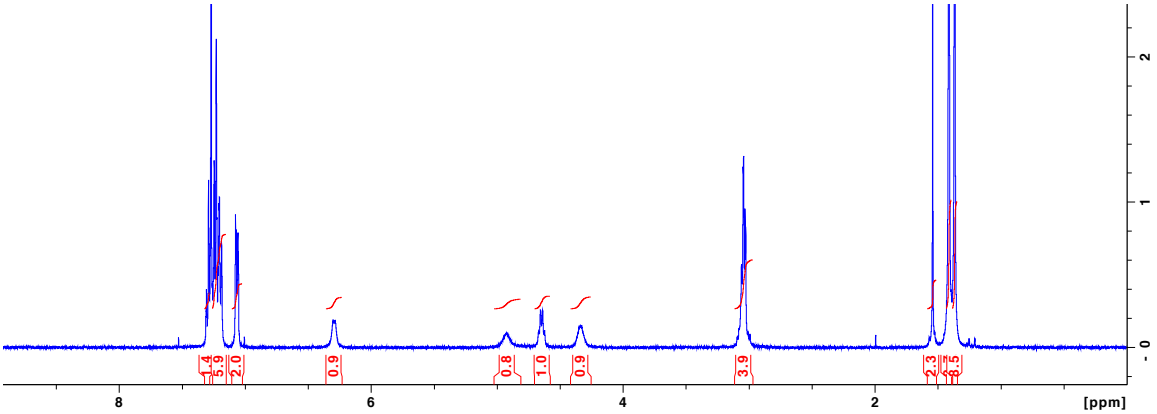
A representative synthesis of the LD-**7la** is described as follows: Boc-L-leucine (2.47 g, 9.9 mmol) and HOBt (1.47 g, 10.9 mmol) were dissolved in dichloromethane (100 mL) at 0 ºC. EDCI (2.09 g, 10.9 mmol) was added subsequently and the solution was stirred at 0 ºC for 30 min. *N*-methyl morpholine (3.27 mL, 29.7 mmol) and D-alanine benzyl ester (1.77 g, 9.9 mmol) were added. The reaction was stirred at ambient temperature for 12 h under Ar. The resulting solution was washed with 10 % citric acid, saturated $NaHCO_3$ solution and brine. The organic layer was dried over $MgSO_4$ and filtered. The solvent was removed under vacuum to obtain 3.80 g (98 %) of white solid, which was used in next step without further purification.

***tert*-Butyl (*tert*-butoxycarbonyl)-*L*-phenylalanyl-*L*-phenylalaninate (LL-7ff)**

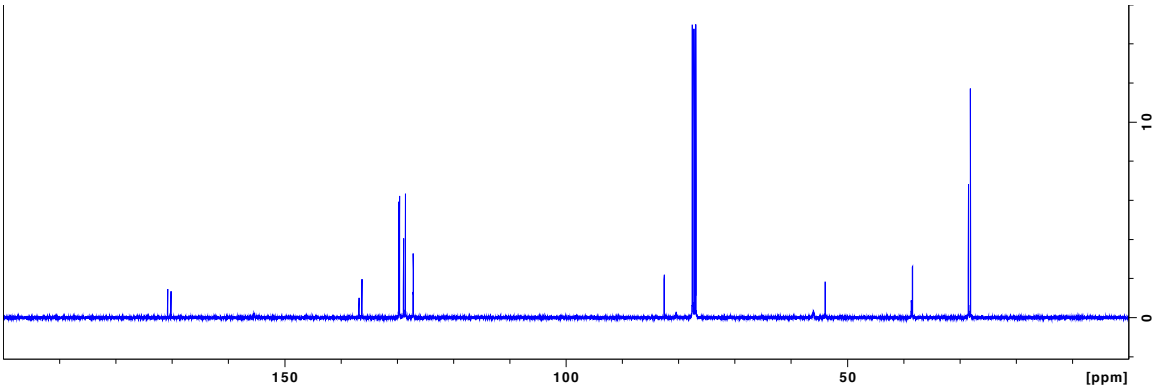


Data for LL-7ff. White solid, 87 % yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.31-7.05 (m, 8H), 7.06 (m, 2H), 6.30 (br, 1H), 4.92 (br, 1H), 4.65 (q, *J* = 6.4 Hz, 1H), 4.33 (br, 1H), 3.05 (m, 4H), 1.41 (s, 9H), 1.37 (s, 9H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 170.8, 170.2, 136.8, 136.3, 129.7, 129.6, 128.9, 128.5, 127.2, 127.1, 82.5, 53.9, 38.4, 28.5, 28.1; MS (ESI) m/z calcd for $C_{27}H_{36}N_2O_5Na^+$ 491.25; found 491.25 $(M+Na)^+$.

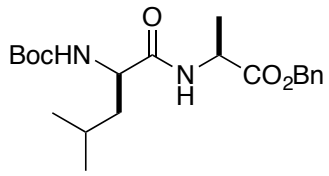
^{1}H NMR spectrum:



^{13}C NMR spectrum:

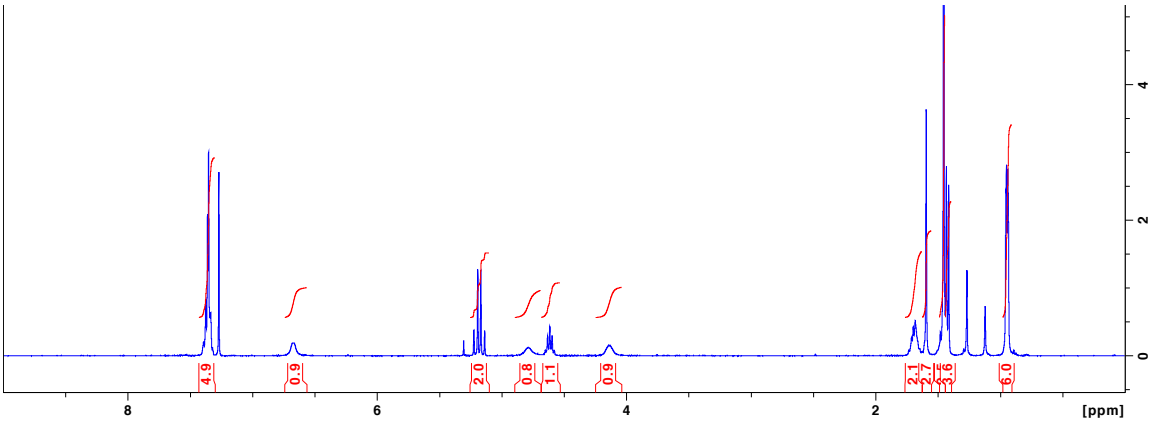


Benzyl (*tert*-butoxycarbonyl)-*D*-leucyl-*L*-alaninate (DL-7la)

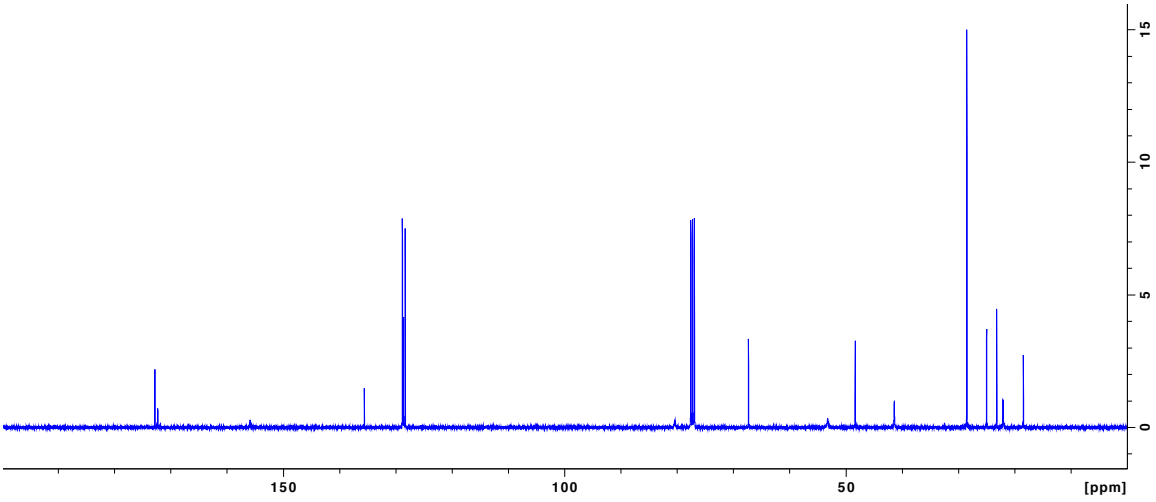


Data for DL-7la. White solid, 88 % yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.36 (m, 5H), 6.67 (br, 1H), 5.18 (dd, *J* = 22.2, 12.3 Hz, 2H), 4.79 (br, 1H), 4.61 (quint, *J* = 7.1 Hz, 1H), 4.13 (br, 1H), 1.70 (m, 2H), 1.45 (s, 9H), 1.44 (m, 1H), 1.42 (d, *J* = 7.1 Hz, 3H), 0.95 (d, *J* = 4.2 Hz, 6H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 172.8, 172.3, 156.0, 135.6, 128.8, 128.6, 128.3, 80.4, 67.3, 53.2, 48.3, 41.4, 28.5, 25.0, 23.2, 22.1, 18.4; MS (ESI) m/z calcd for $C_{21}H_{32}N_2O_5Na^+$ 415.22; found 415.22 $(M+Na)^+$.

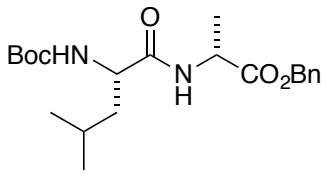
^{1}H NMR spectrum:



^{13}C NMR spectrum:

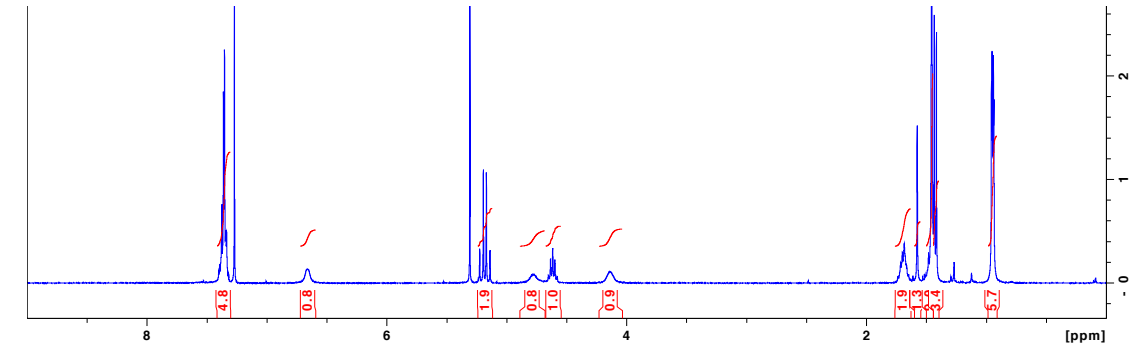


Benzyl (*tert*-butoxycarbonyl)-*L*-leucyl-*D*-alaninate (LD-7la)

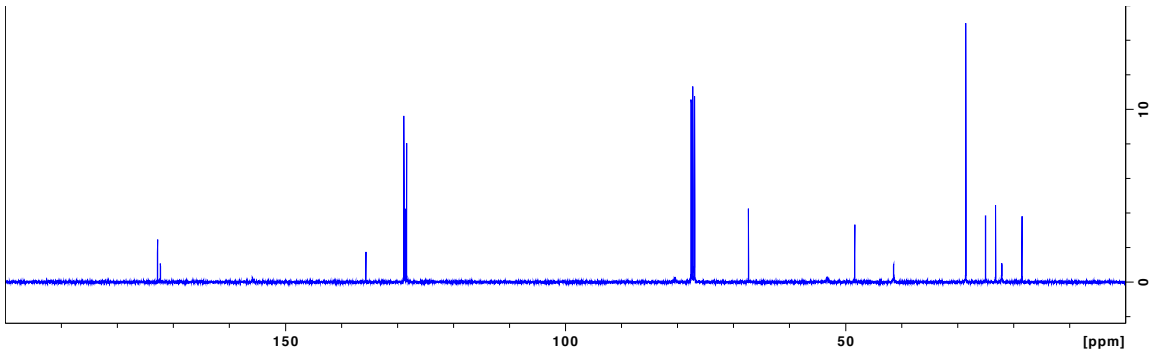


Data for LD-7la. White solid, 99 % yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.38-7.35 (m, 5H), 6.66 (br, 1H), 5.18 (dd, *J* = 22.2, 12.3 Hz, 2H), 4.78 (br, 1H), 4.61 (quint, *J* = 7.1 Hz, 1H), 4.13 (br, 1H), 1.69 (m, 2H), 1.45 (s, 9H), 1.44 (m, 1H), 1.42 (d, *J* = 7.1 Hz, 3H), 0.95 (d, *J* = 4.2 Hz, 6H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 172.8, 172.3, 156.0, 135.6, 128.8, 128.6, 128.3, 80.4, 67.3, 53.2, 48.3, 41.4, 28.5, 25.0, 23.2, 22.1, 18.5; MS (ESI) m/z calcd for $C_{21}H_{32}N_2O_5Na^+$ 415.22; found 415.22 $(M+Na)^+$.

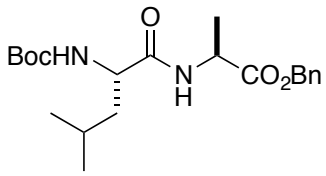
^{1}H NMR spectrum:



^{13}C NMR spectrum:

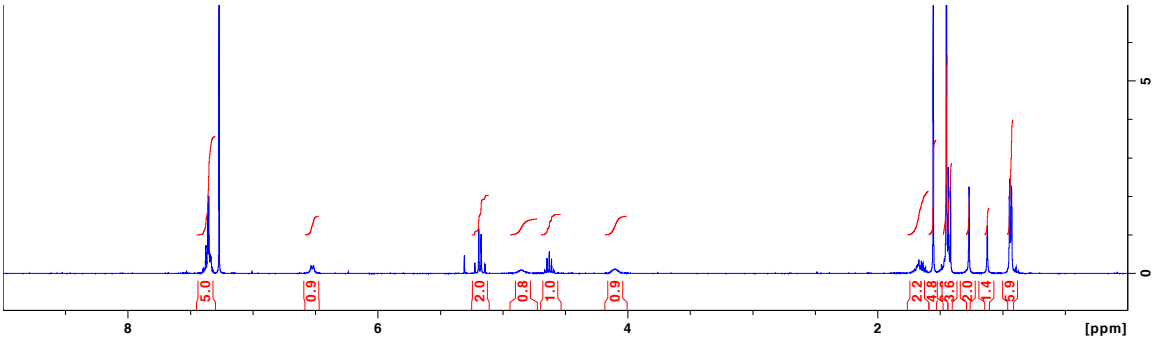


Benzyl (*tert*-butoxycarbonyl)-*L*-leucyl-*L*-alaninate (LL-7la)

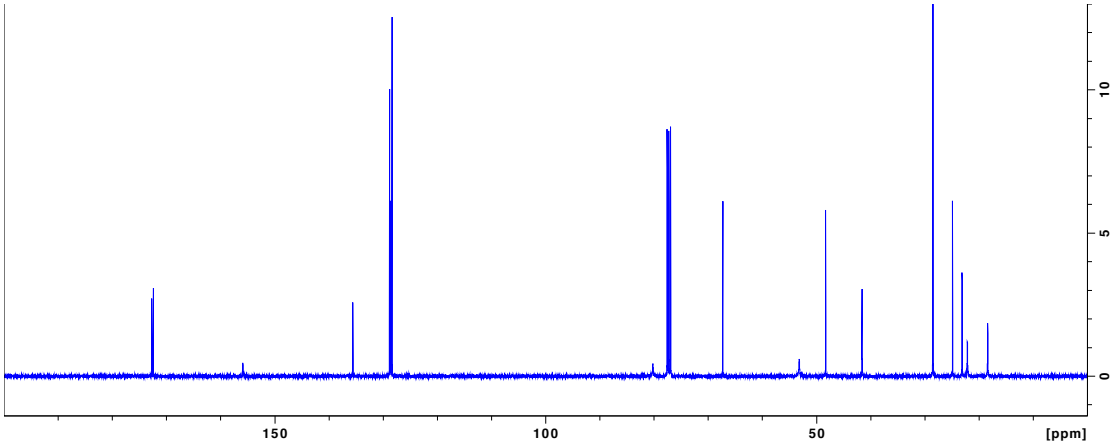


Data for LL-7la. White solid, 91 % yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.38-7.35 (m, 5H), 6.52 (br, 1H), 5.18 (dd, *J* = 22.2, 12.3 Hz, 2H), 4.84 (br, 1H), 4.63 (quint, *J* = 7.1 Hz, 1H), 4.10 (br, 1H), 1.70-1.62 (m, 2H), 1.58 (s, 1H), 1.45 (s, 9H), 1.44 (m, 1H), 1.43 (d, *J* = 7.1 Hz, 3H), 0.94 (d, *J* = 4.2 Hz, 6H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 172.7, 172.4, 155.9, 135.6, 128.8, 128.6, 128.3, 80.2, 67.3, 53.2, 48.3, 41.6, 28.5, 24.9, 23.1, 22.1, 18.4; MS (ESI) m/z calcd for $C_{21}H_{32}N_2O_5Li^+$ 399.25; found 399.25 $(M+Li)^+$

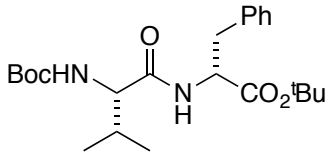
^{1}H NMR spectrum:



^{13}C NMR spectrum:

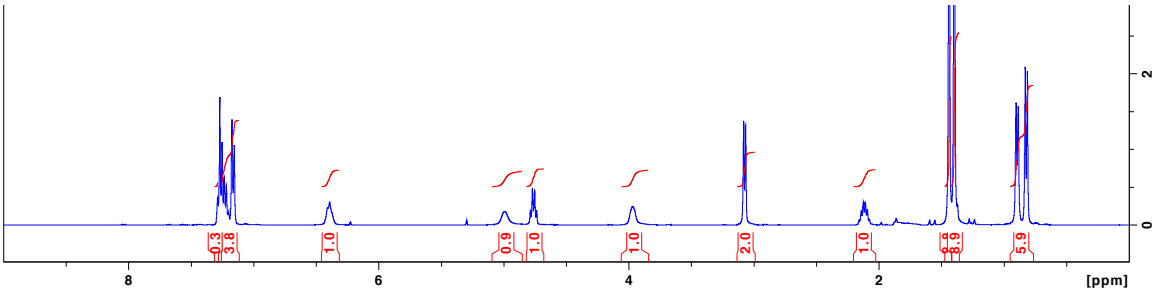


***tert*-Butyl (*tert*-butoxycarbonyl)-*L*-valyl-*D*-phenylalaninate (LD-7vf)**

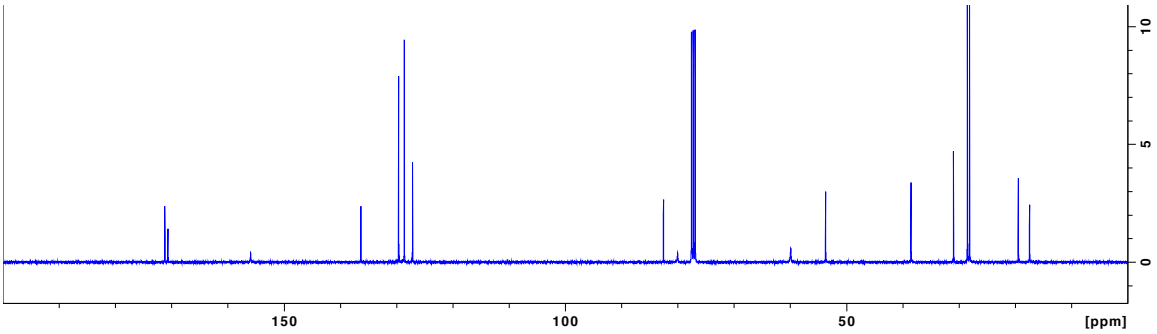


Data for LD-7vf. White solid, 90 % yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.29-7.16 (m, 5H), 6.39 (br, 1H), 4.99 (br, 1H), 4.76 (dd, *J* = 13.9, 6.4 Hz, 1H), 3.97 (br 1H), 3.08 (d, *J* = 6.2 Hz, 2H), 2.1 (m, 1H), 1.44 (s, 9H), 1.40 (s, 9H), 0.90 (d, *J* = 6.7 Hz, 3H), 0.82 (d, *J* = 6.7 Hz, 3H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 171.2, 170.6, 155.9, 136.3, 129.6, 128.6, 127.2, 82.5, 80.0, 59.9, 53.7, 38.5, 31.0, 28.5, 28.1, 19.4, 17.4; MS (ESI) m/z calcd for $C_{23}H_{36}N_2O_5Na^+$ 443.25; found 443.26 $(M+Na)^+$.

^{1}H NMR spectrum:



^{13}C NMR spectrum:

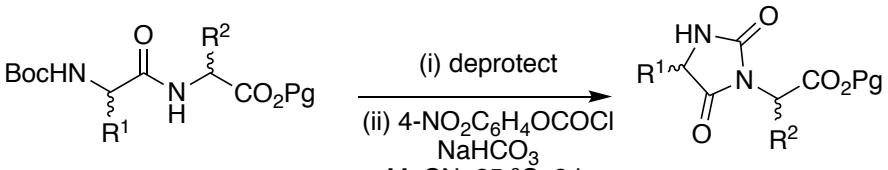


General Procedure of Boc Deprotection

Method A: This method was used for LL-**7ff** and LD-**7vf**. The solution of 4 M HCl in dioxane (10 mL) was cooled down to 0 ºC in ice bath under Ar. Then, this solution was transferred to pre-cool flask containing Boc-dipeptide (0.50 mmol). The solution was stirred at ambient temperature for additional 30 min until starting material was disappeared as monitored by TLC. The solvent was removed under vacuum. The crude oil was triturated with ether and filtered to obtain the product as a salt.

Method B: This method was used for DL-**7la**, LD-**7la** and LL-**7la**. Boc-dipeptide (1 mmol) was dissolved in 3 mL of dichloromethane, then TFA (1 mL) was added and the solution was stirred for 1 h. The solution was completely evaporated by rotavap, redissolved in dichloromethane and evaporated for at least 3 times to remove excess TFA. The crude oil was used in the next step without further purification.

*General Procedure for the Preparation of C-protected Hydantoin **8***



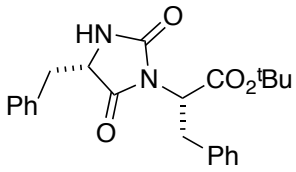
7 **8**

LL-**8ff**, 72%; DL-**8la**, 44%
LD-**8la**, 39%; LL-**8la**, 58%
LD-**8vf**, 57%

A representative synthesis of the LL-**8ff** is described as follows: Boc-deprotected LL-**7ff** (1.35 g, 3 mmol), $NaHCO_3$ (1.26 g, 15 mmol) were dissolved in acetonitrile (48 mL). Then, *p*-nitrophenyl chloroformate (0.79 g, 3.9 mmol) was added. The cloudy solution was stirred at room temperature for 3 h under Ar. Et_3N (0.54 mL) was added, and the

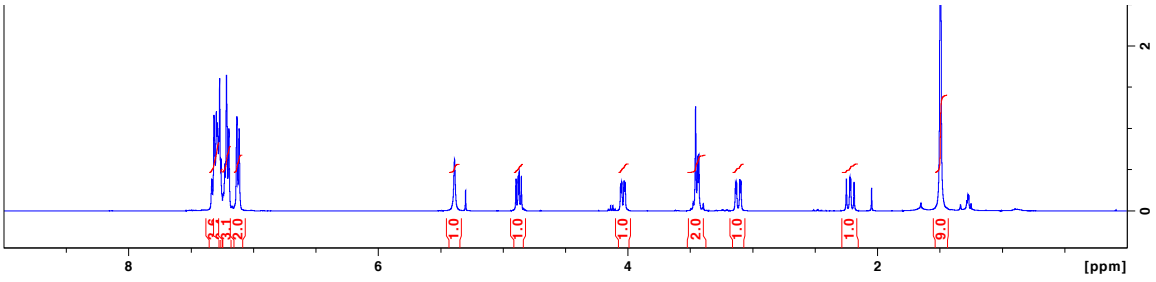
yellow solution was stirred for additional 12 h. After removal of acetonitrile under vacuum, the crude was dissolved in ethyl acetate, washed 5 times with 5% K_2CO_3, brine and dried over $MgSO_4$. The organic solution was removed to obtain the crude oil. The residue was purified by column chromatography on silica gel with hexanes-ethyl acetate (3:1) to afford the product (0.85 g, 72%) as pale yellow solid.

***tert*-Butyl (*S*)-2-((*S*)-4-benzyl-2,5-dioxoimidazolidin-1-yl)-3-phenylpropanoate (LL-8ff)**

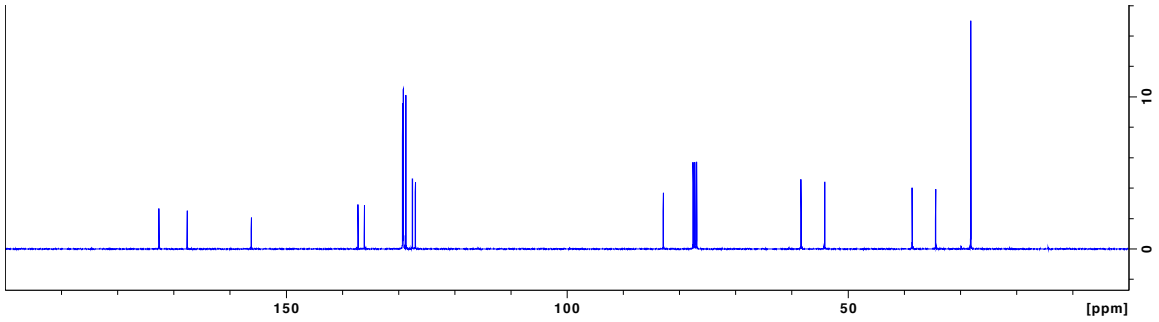


Data for LL-8ff. pale yellow solid, 72 % yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.33-7.11 (m, 10H), 5.39 (br, 1H), 4.88 (dd, *J* = 10.1, 6.9 Hz, 1H), 4.04 (dd, *J* = 10.7, 3.3 Hz, 1H), 3.46-3.43 (m, 2H), 3.12 (dd, *J* = 13.9, 3.5 Hz, 1H), 2.22 (dd, *J* = 13.8, 10.9 Hz, 1H), 1.49 (s, 1H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 172.6, 167.6, 156.2, 137.2, 136.1, 129.3, 129.2, 129.1, 128.7, 127.5, 127.0, 82.9, 58.3, 54.1, 38.6, 34.4, 28.1; HRMS (ESI) m/z calcd for $C_{23}H_{36}N_2O_4Na^+$ 417.1785; found 417.1807 $(M+Na)^+$

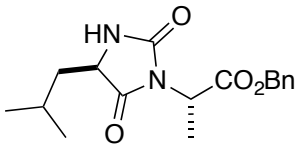
^{1}H NMR spectrum:



^{13}C NMR spectrum:

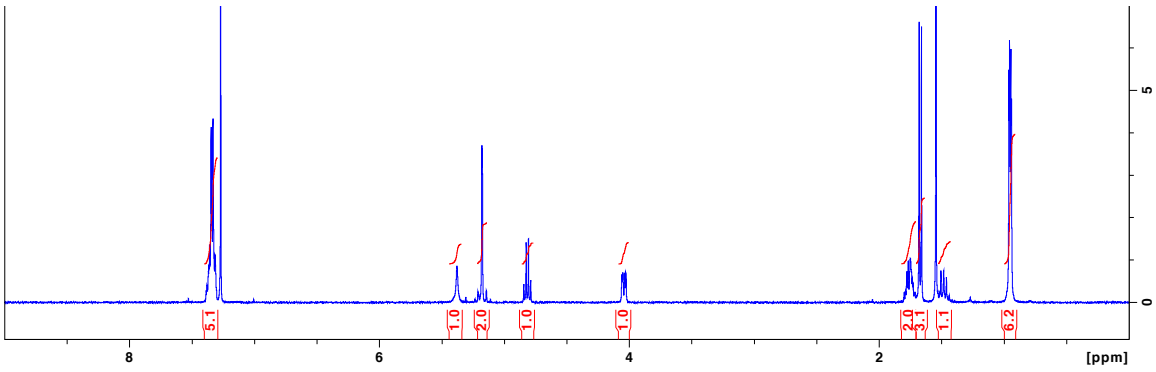


Benzyl (*S*)-2-((*R*)-4-isobutyl-2,5-dioxoimidazolidin-1-yl)propanoate (DL-8la)

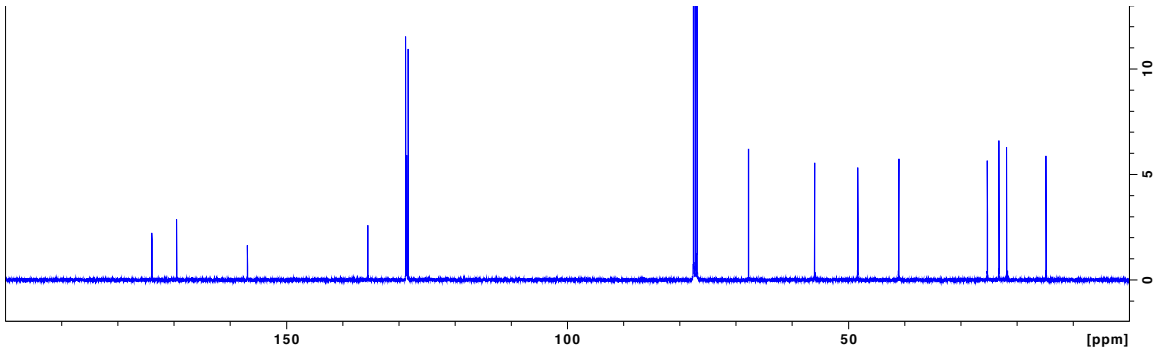


Data for DL-8la. White solid, 44 % yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.40-7.30 (m, 5H), 5.38 (br 1H), 5.22-5.14 (m, 2H), 4.82 (q, *J* = 7.4, 1H), 4.05 (m, 1H), 1.80 (m, 2H), 1.67 (d, *J* = 7.4, 3H), 1.50-1.46 (m, 1H), 0.95 (m, 6H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 173.9, 169.5, 156.9, 135.5, 128.8, 128.6, 128.3, 67.7, 56.0, 48.3, 41.0, 25.2, 23.2, 21.8, 14.8; HRMS (ESI) m/z calcd for $C_{17}H_{23}N_2O_4^+$ 319.1652; found 319.1646 $(M+H)^+$

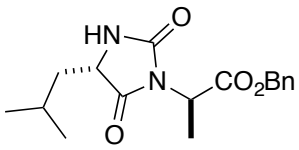
^{1}H NMR spectrum:



^{13}C NMR spectrum:

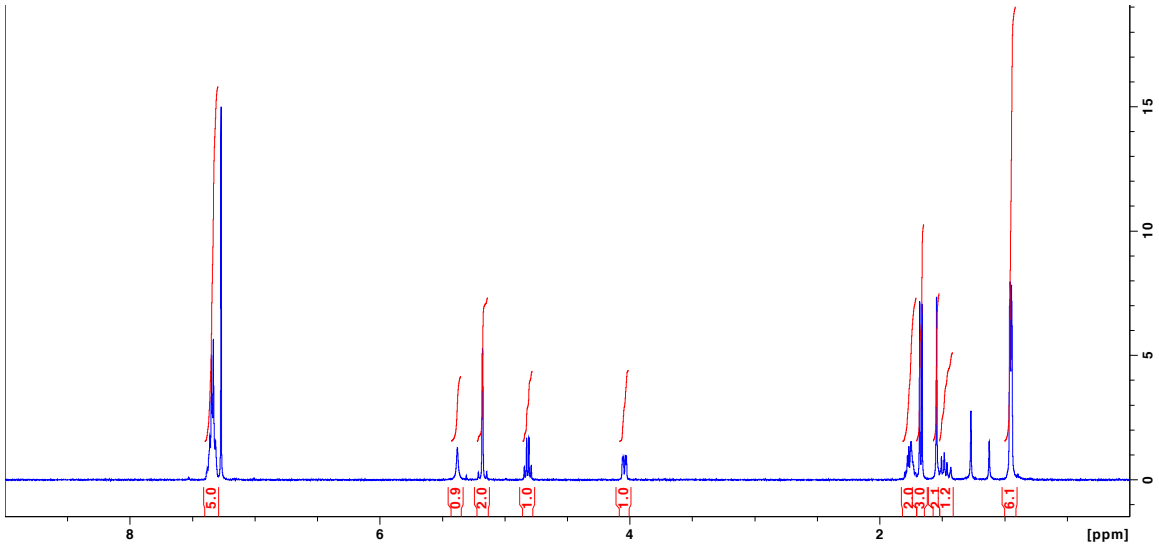


Benzyl (*R*)-2-((*S*)-4-isobutyl-2,5-dioxoimidazolidin-1-yl)propanoate (LD-8la)

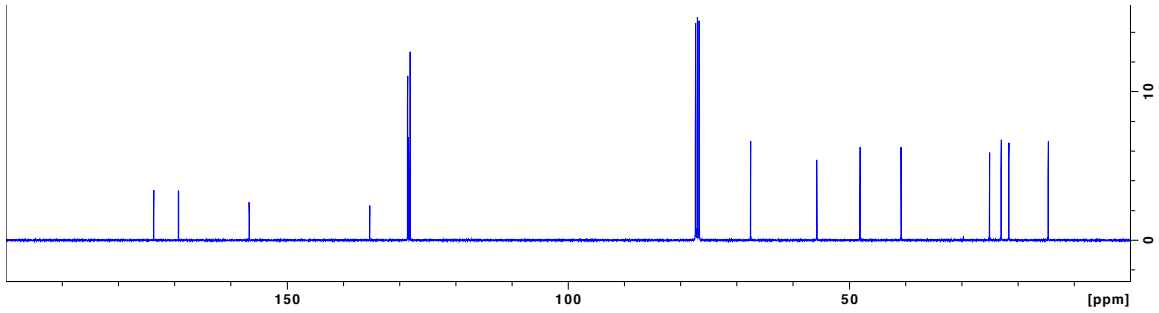


Data for LD-8la. White solid, 39 % yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.40-7.30 (m, 5H), 5.38 (br 1H), 5.22-5.14 (m, 2H), 4.82 (q, *J* = 7.4, 1H), 4.05 (m, 1H), 1.80 (m, 2H), 1.67 (d, *J* = 7.4, 3H), 1.54-1.46 (m, 1H), 0.95 (m, 6H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 173.7, 169.3, 156.7, 135.3, 128.5, 128.4, 128.1, 67.5, 55.8, 48.1, 40.8, 25.0, 23.0, 21.6, 14.6; HRMS (ESI) m/z calcd for $C_{17}H_{22}N_2O_4Na^+$ 341.1472; found 341.1491 $(M+Na)^+$

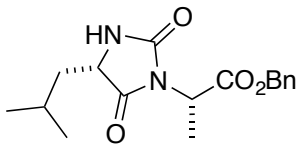
^{1}H NMR spectrum:



^{13}C NMR spectrum:

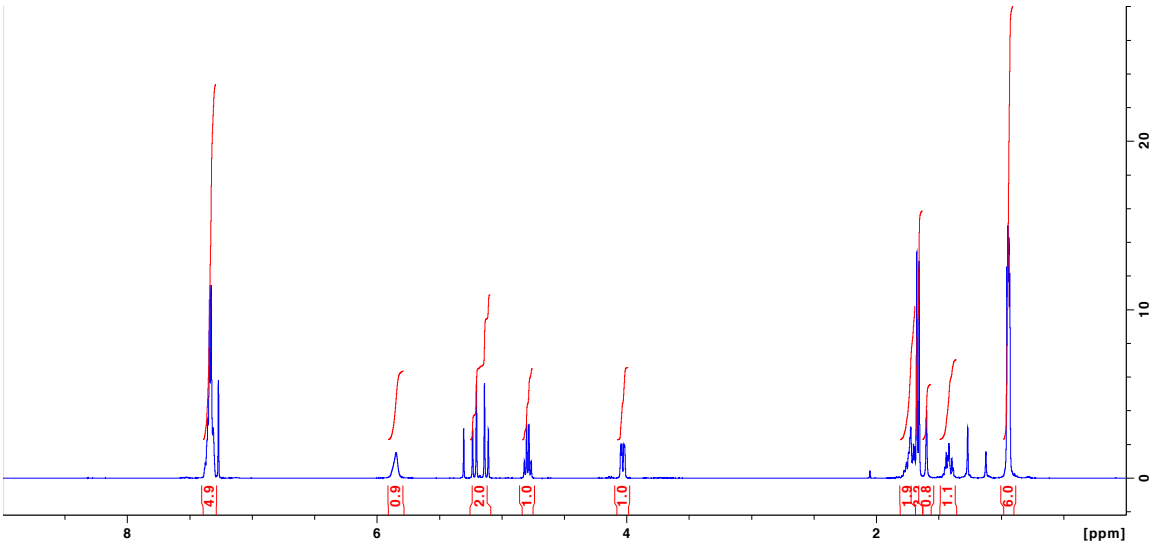


Benzyl (*S*)-2-((*S*)-4-isobutyl-2,5-dioxoimidazolidin-1-yl)propanoate (LL-8la)

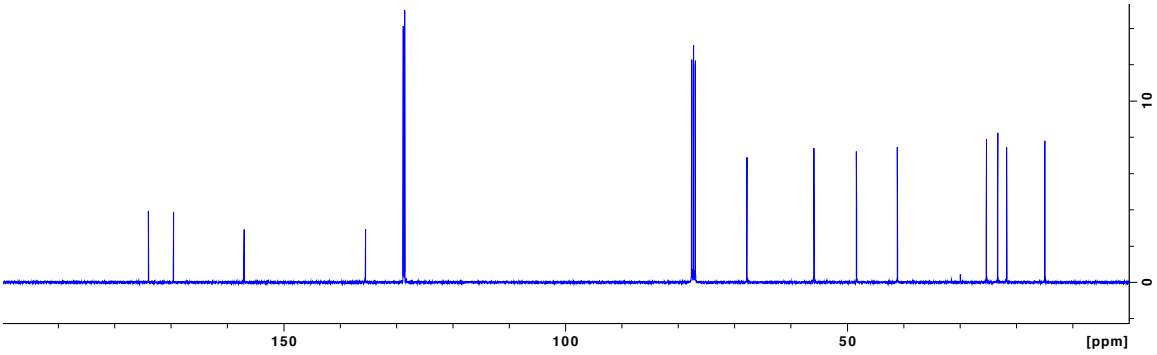


Data for LL-8la. White solid, 39 % yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.40-7.30 (m, 5H), 5.85 (br, 1H), 5.17 (dd, *J* = 38.1, 12.3 Hz, 2H), 4.79 (q, *J* = 7.3 Hz, 1H), 4.06-4.01 (m, 1H), 1.81-1.69 (m, 2H), 1.67 (d, *J* = 7.4 Hz, 3H), 1.50-1.36 (m, 1H), 0.95 (m, 6H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 174.0, 169.5, 157.0, 135.5, 128.8, 128.6, 128.5, 67.8, 55.9, 48.3, 41.1, 25.3, 23.3, 21.7, 14.9; HRMS (ESI) m/z calcd for $C_{17}H_{23}N_2O_4^+$ 319.1652; found 319.1642 $(M+H)^+$

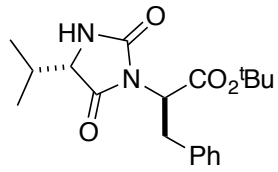
^{1}H NMR spectrum:



^{13}C NMR spectrum:

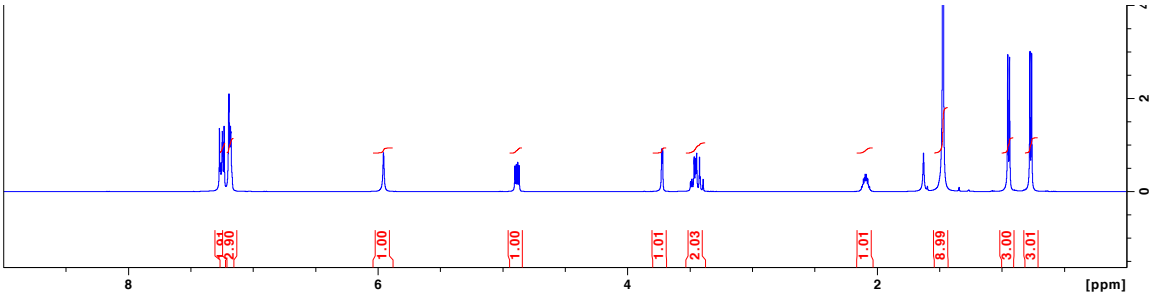


***tert*-Butyl (*R*)-2-((*S*)-4-isopropyl-2,5-dioxoimidazolidin-1-yl)-3-phenylpropanoate (LD-8vf)**

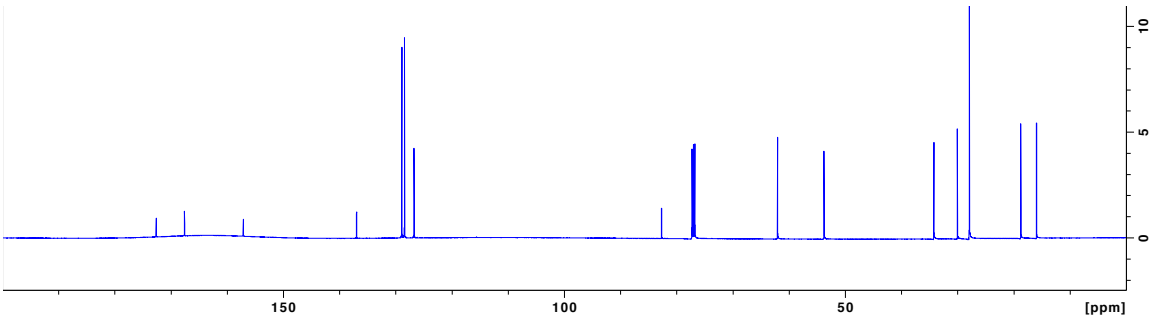


Data for LD-8vf. solid, 57 % yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.29-7.16 (m, 5H), 5.96 (br, 1H), 4.89 (dd, *J* = 11.3, 5.5 Hz, 1H), 3.72 (dd, *J* = 3.8, 1.3 Hz, 1H), 3.5-3.39 (m, 2H), 2.16-2.02 (m, 1H), 1.47 (s, 9H), 0.95 (d, *J* = 7.0 Hz, 3H), 0.78 (d, *J* = 6.8 Hz, 3H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 172.6, 167.6, 157.1, 137.0, 128.9, 128.4, 126.7, 82.7, 62.0, 53.8, 34.2, 30.0, 27.9, 18.7, 16.0; HRMS (ESI) m/z calcd for $C_{19}H_{26}N_2O_4Na^+$ 369.1785; found 369.1814 $(M+Na)^+$

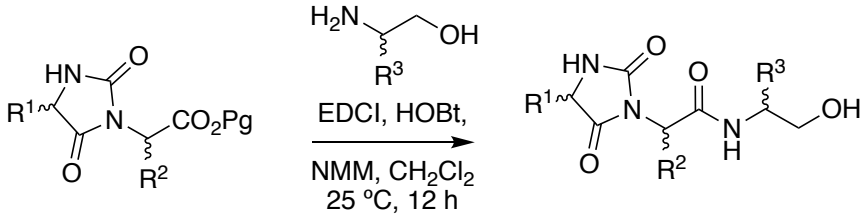
^{1}H NMR spectrum:



^{13}C NMR spectrum:



*Procedure for the Preparation of Scaffold **9***



8

9

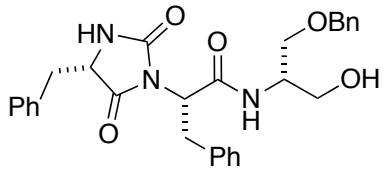
LLD-**ffs'**, 68%; LLL-**fff**, 83%
DLD-**lal**, 88%; LDL-**lal**, 63%
LLD-**lal**, 58%; LDL-**vfc'**, 85%

Method A: (*tert*-butyl ester protecting group): A representative synthesis of the LLL-**9fff** is described as follows: Compound LL-**8ff** (0.2 g, 0.51 mmol) was dissolved in dichloromethane (4 mL). TFA (1.3 mL) was added and the solution was stirred at room temperature for 3 h. The solvent was removed under vacuum, and then repeated dissolve/evaporate cycles with dichloromethane several times to remove excess TFA. The crude oil was used for the next step without further purification. The *C*-deprotected hydantoin, HOBt (76 mg, 0.56 mmol) were dissolved in dichloromethane (5 mL) in an ice bath. EDCI (108 mg, 0.56 mmol) was added and the solution was stirred for 30 min at 0 ºC. *N*-methyl morpholine (0.17 mL, 1.53 mmol) and L-phenylalaninol (77 mg, 0.51 mmol) were added subsequently. The solution was stirred at ambient temperature for 12 h under Ar. Organic layer was washed with 1 N HCl solution, saturated $NaHCO_3$, brine and dried over $MgSO_4$ successively. The solution was removed under reduced pressure to obtain the product as white solid (199 mg, 83 %).

Method B: (benzyl ester protecting group): A representative synthesis of the LLD-**9lal** is described as follows: Compound LL-**8la** (0.47 g, 1.49 mmol) and 10 % Pd/C (0.15 g, 0.15 mmol) were dissolved in dried methanol (20 mL) under Ar. Then, H_2 was purged

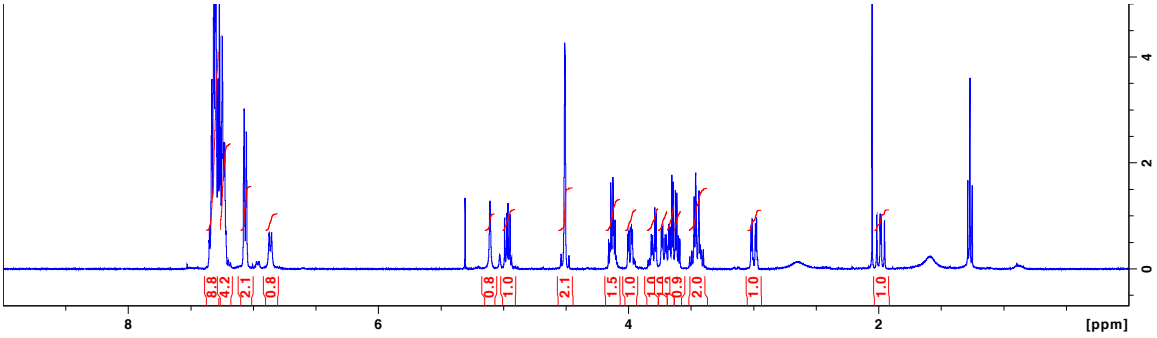
into the flask for 5 min and the solution was stirred under H_2 atmosphere for 12 h. The solution was filtered through celite. The clear solution was removed under vacuum to obtain the oil product (0.34 g, 100 %), which was used in the next step without further purification. The *C*-deprotected hydantoin (0.34 g, 1.48 mmol), HOBt (0.26 g, 1.93 mmol) were dissolved in dichloromethane (20 mL) in an ice bath. EDCI (0.37 g, 1.93 mmol) was added and the solution was stirred for 30 min at 0 ºC. *N*-methyl morpholine (0.49 mL, 4.45 mmol) and D-leucinol (0.22 g, 1.93 mmol) were added subsequently. The solution was stirred at ambient temperature for 12 h under Ar. Organic layer was washed with 1 N HCl solution, saturated $NaHCO_3$, brine and dried over $MgSO_4$ successively. The solution was removed under reduced pressure to obtain white solid (0.28 g, 58 %).

(*S*)-2-((*S*)-4-Benzyl-2,5-dioxoimidazolidin-1-yl)-*N*-((*S*)-1-(benzyloxy)-3-hydroxypropan-2-yl)-3-phenylpropanamide (LLD-9ffs')

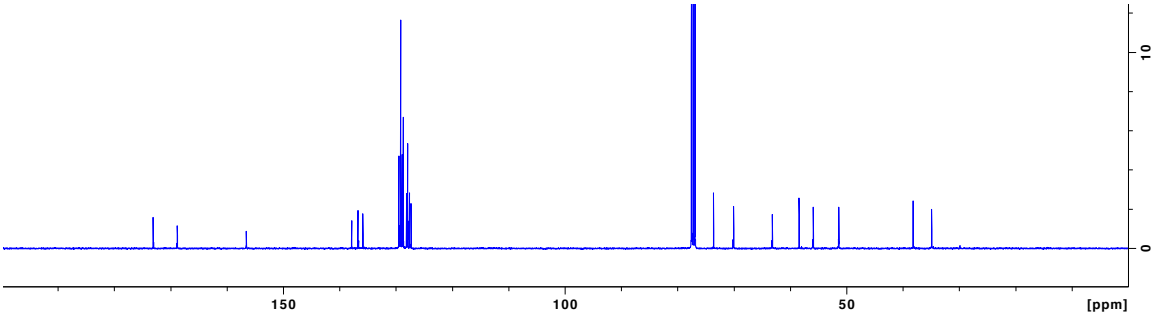


Data for LLD-9ffs'. Solid, 68 % yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.37-7.18 (m, 13 H), 7.10-7.01 (m, 2H), 6.86 (d, *J* = 7.6 Hz, 1H), 5.11 (s, 1H), 4.97 (dd, *J* = 10.7, 6.5 Hz, 1H), 4.51 (s, 2H), 4.18-4.07 (m, 2H), 4.03-3.93 (m, 1H), 3.85-3.69 (m, 2H), 3.69-3.58 (m, 2H), 3.52-3.37 (m, 2H), 3.00 (dd, *J* = 13.8, 3.6 Hz, 1H), 1.99 (dd, *J* = 13.9, 10.8 Hz, 1H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 173.1, 168.8, 156.6, 137.8, 136.7, 135.9, 129.4, 129.2, 129.2, 128.9, 128.7, 128.1, 127.9, 127.6, 127.3, 73.6, 70.0, 63.2, 58.4, 55.9, 51.4, 38.2, 34.9; HRMS (ESI) m/z calcd for $C_{29}H_{30}N_3O_5^-$ 500.2191; found 500.2174 $(M-H)^-$

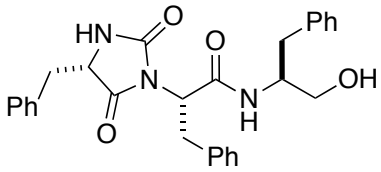
^{1}H NMR spectrum:



^{13}C NMR spectrum:

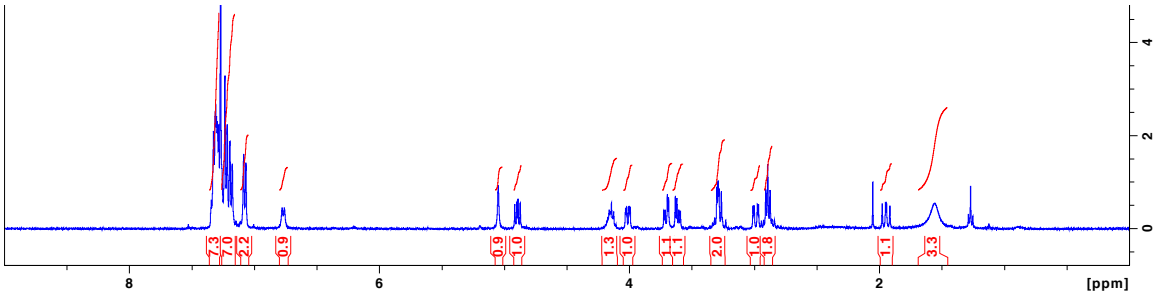


(*S*)-2-((*S*)-4-Benzyl-2,5-dioxoimidazolidin-1-yl)-*N*-((*S*)-1-hydroxy-3-phenylpropan-2-yl)-3-phenylpropanamide (LLL-9fff)

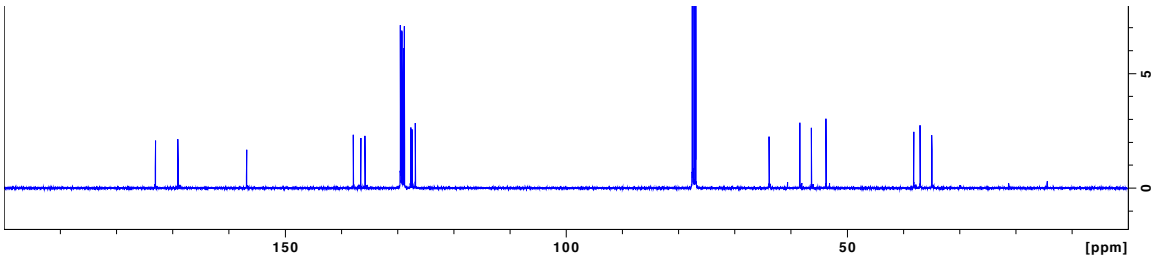


Data for LLL-9fff. Solid, 83 % yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.37-7.15 (m, 14H), 7.08 (m, 2H), 6.79 (d, *J* = 7.2 Hz, 1H), 5.05 (s, 1H), 4.90 (dd, *J* = 10.5, 6.8 Hz, 1H), 4.22-4.10 (m, 1H), 4.05-3.98 (m, 1H), 3.71 (dd, *J* = 11.3, 3.6 Hz, 1H), 3.61 (dd, *J* = 11.2, 5.2 Hz, 1H), 3.35-3.23 (m, 2H), 2.99 (dd, *J* = 13.8, 3.6 Hz, 1H), 2.92-2.86 (m, 2H), 1.95 (dd, *J* = 13.8, 10.7 Hz, 1H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 173.1, 169.0, 156.8, 137.9, 136.5, 135.8, 129.5, 129.4, 129.2, 129.1, 128.9, 128.8, 127.6, 127.4, 126.8, 63.9, 58.4, 56.4, 53.8, 38.1, 37.0, 34.9; HRMS (ESI) m/z calcd for $C_{28}H_{28}N_3O_4^-$ 470.2085; found 470.2062 $(M-H)^-$

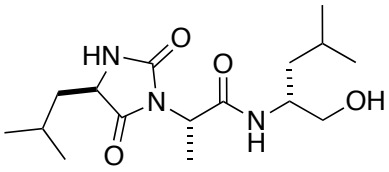
^{1}H NMR spectrum:



^{13}C NMR spectrum:

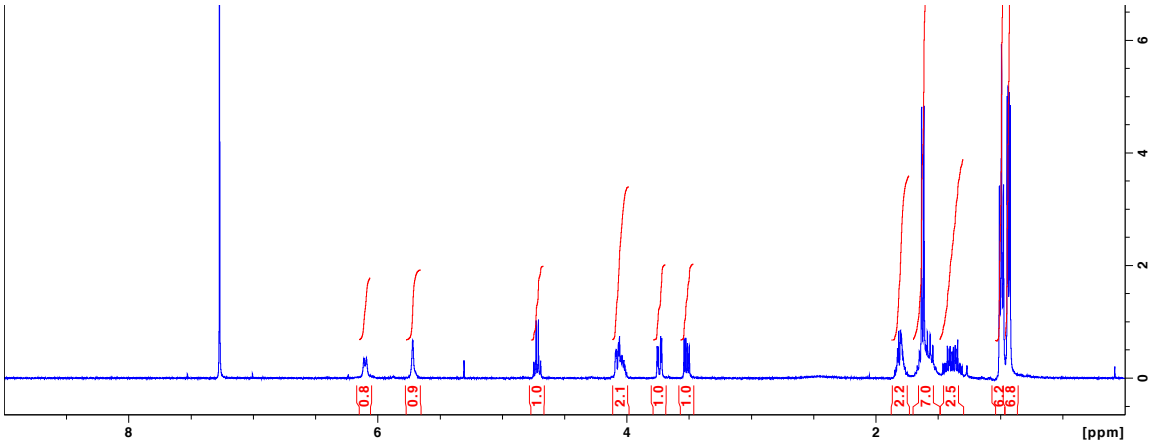


(*S*)-*N*-((*R*)-1-Hydroxy-4-methylpentan-2-yl)-2-((*R*)-4-isobutyl-2,5-dioxoimidazolidin-1-yl)propanamide (DLD-9lal)

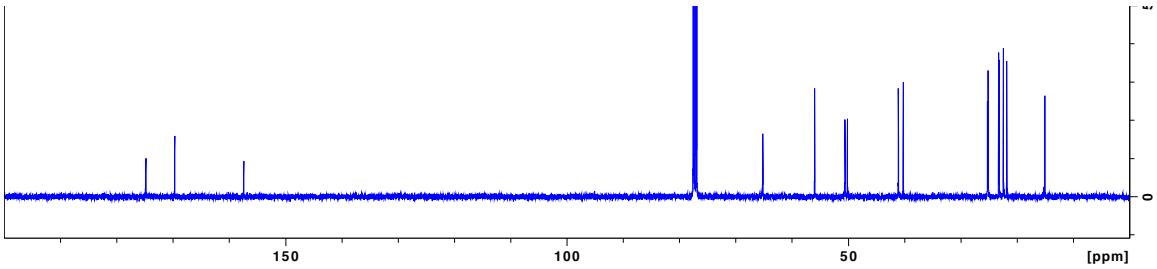


Data for DLD-9lal. White solid, 88 % yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 6.10 (br, 1H), 5.72 (br, 1H), 4.72 (q, *J* = 7.3 Hz, 1H), 4.11-3.98 (m, 2H), 3.74 (dd, *J* = 11.2, 3.5 Hz, 1H), 3.52 (dd, *J* = 11.3, 5.5 Hz, 1H), 1.88-1.73 (m, 2H), 1.70-1.49 (m, 6H), 1.49-1.30 (m, 2H), 0.99 (t, *J* = 6.6 Hz, 6H), 0.94 (dd, *J* = 6.6, 2.9 Hz, 6H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 174.8, 169.7, 157.4, 65.2, 56.0, 50.6, 50.1, 41.1, 40.2, 25.2, 25.1, 23.3, 23.2, 22.4, 21.8, 15.1; HRMS (ESI) m/z calcd for $C_{16}H_{30}N_3O_4^+$ 328.2231; found 328.2253 $(M+H)^+$

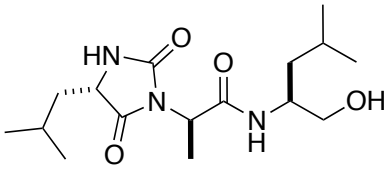
^{1}H NMR spectrum:



^{13}C NMR spectrum:

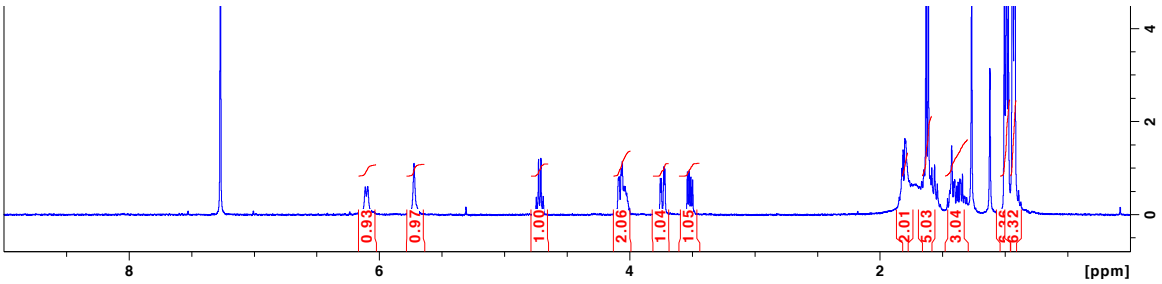


***R*)-*N*-((*S*)-1-Hydroxy-4-methylpentan-2-yl)-2-((*S*)-4-isobutyl-2,5-dioxoimidazolidin-1-yl)propanamide (LDL-9lal)**

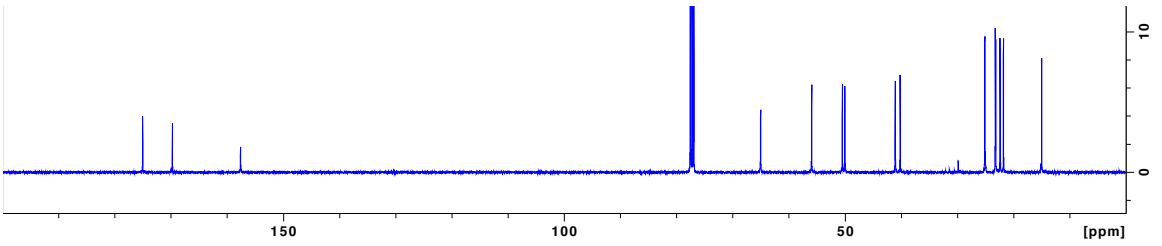


Data for LDL-9lal. White solid, 63 % yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 6.10 (br, 1H), 5.72 (br, 1H), 4.72 (q, *J* = 7.3 Hz, 1H), 4.11-3.98 (m, 2H), 3.74 (dd, *J* = 11.3, 3.3 Hz, 1H), 3.52 (dd, *J* = 11.3, 5.5 Hz, 1H), 1.88-1.76 (m, 2H), 1.66-1.58 (m, 5H), 1.49-1.30 (m, 3H), 0.99 (t, *J* = 6.6 Hz, 6H), 0.94 (dd, *J* = 6.6, 2.9 Hz, 6H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 175.0, 169.7, 157.6, 65.0, 55.9, 50.5, 50.0, 41.1, 40.2, 25.2, 25.1, 23.3, 23.2, 22.4, 21.8, 15.0; HRMS (ESI) m/z calcd for $C_{16}H_{30}N_3O_4^+$ 328.2231; found 328.2228 $(M+H)^+$

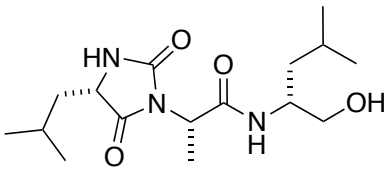
^{1}H NMR spectrum:



^{13}C NMR spectrum:

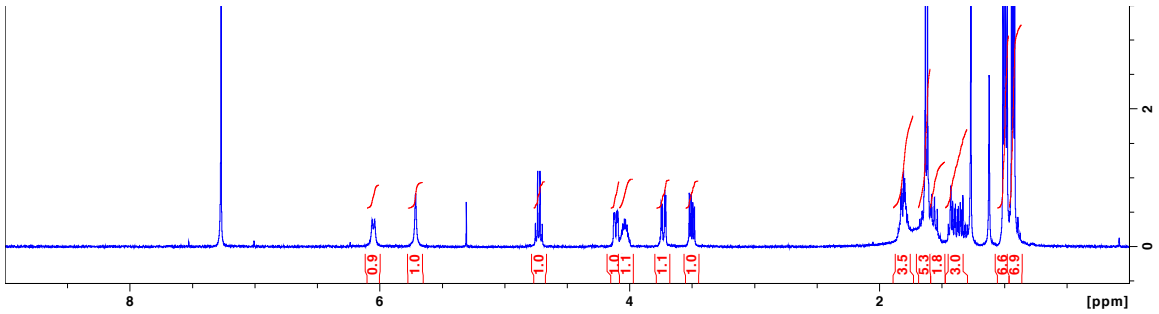


(*S*)-*N*-((*R*)-1-Hydroxy-4-methylpentan-2-yl)-2-((*S*)-4-isobutyl-2,5-dioxoimidazolidin-1-yl)propanamide (LLD-9lal)

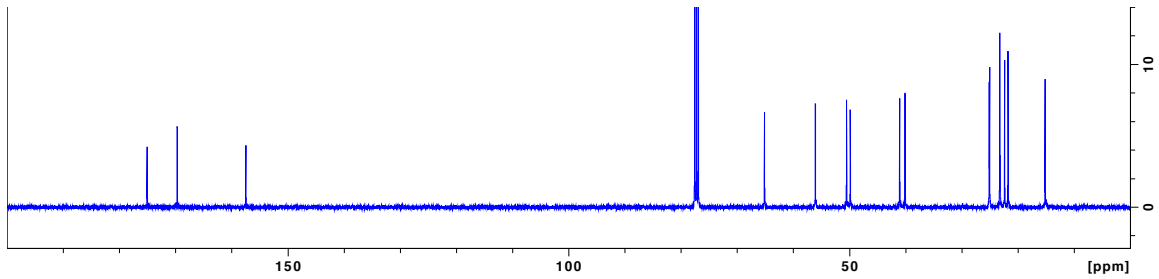


Data for LLD-9lal. White solid, 58 % yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 6.05 (br, 1H), 5.71 (br, 1H), 4.73 (q, *J* = 7.3 Hz, 1H), 4.11 (m, 1H), 4.08-3.99 (m, 1H), 3.73 (dd, *J* = 11.2, 3.5 Hz, 1H), 3.52 (dd, *J* = 11.3, 5.5 Hz, 1H), 1.86-1.76 (m, 2H), 1.65-1.53 (m, 5H), 1.47-1.29 (m, 3H), 0.99 (t, *J* = 6.6 Hz, 6H), 0.94 (dd, *J* = 6.6, 3.4 Hz, 6H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 175.1, 169.7, 157.5, 65.1, 56.1, 50.5, 49.9, 41.0, 40.1, 25.1, 25.0, 23.2, 23.1, 22.4, 21.8, 15.2; HRMS (ESI) m/z calcd for $C_{16}H_{30}N_3O_4^+$ 328.2231; found 328.2276 $(M+H)^+$

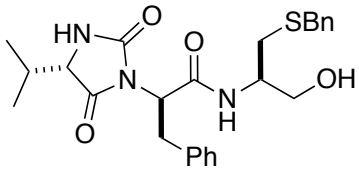
^{1}H NMR spectrum:



^{13}C NMR spectrum:

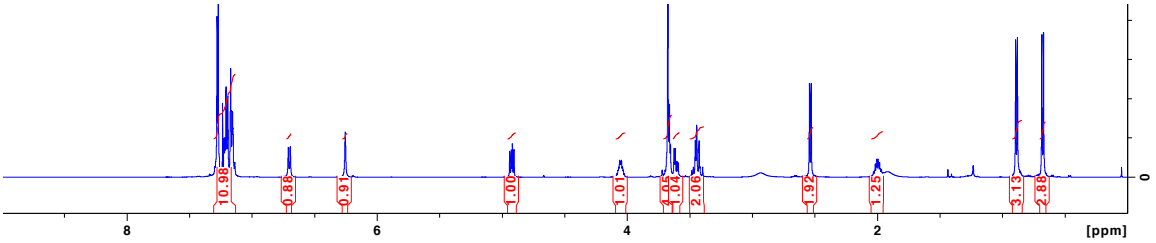


(*R*)-*N*-((*R*)-1-(Benzylthio)-3-hydroxypropan-2-yl)-2-((*S*)-4-isopropyl-2,5-dioxoimidazolidin-1-yl)-3-phenylpropanamide (LDL-9vfc')

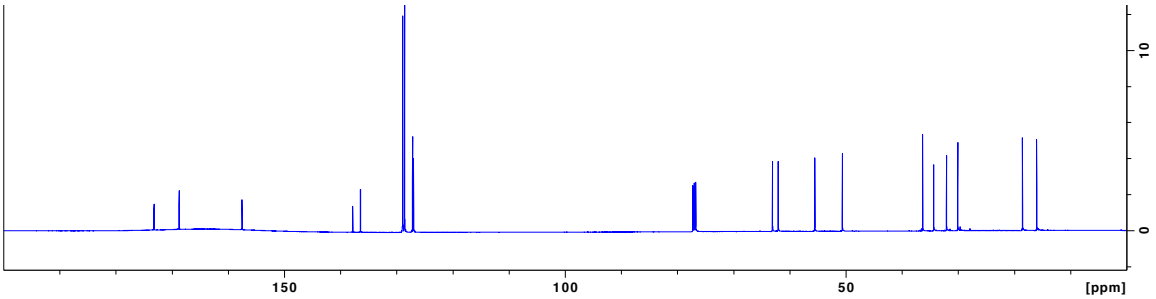


Data for LDL-9vfc'. White solid, 85 % yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.31-7.13 (m, 10H), 6.70 (d, *J* = 7.9 Hz, 1H), 6.25 (s, 1H), 4.92 (dd, *J* = 10.5, 6.5 Hz, 1H), 4.05 (m, 1H), 3.71-3.64 (m, 4H), 3.64-3.58 (m, 1H), 3.50-3.39 (m, 2H), 2.93 (br, 1H), 2.54 (d, *J* = 6.8 Hz, 2H), 2.05-1.95 (m, 1H), 0.89 (d, *J* = 7.0 Hz, 3H), 0.68 (d, *J* = 6.8 Hz, 3H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 173.2, 168.7, 157.5, 137.8, 136.4, 128.9, 128.9, 128.6, 127.2, 127.0, 63.1, 62.1, 55.5, 50.6, 36.3, 34.4, 32.1, 30.1, 18.6, 16.0; HRMS (ESI) m/z calcd for $C_{25}H_{32}N_3O_4S^+$ 470.2108; found 470.2146 $(M+H)^+$

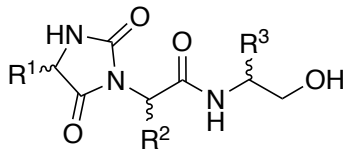
^{1}H NMR spectrum:



^{13}C NMR spectrum:



*General Procedure for the Preparation of Scaffold **1***

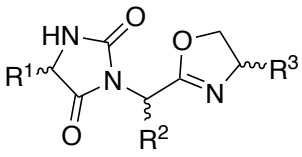


9

LLD-**ffs'**, 68%; LLL-**fff**, 83%
DLD-**lal**, 88%; LDL-**lal**, 63%
LLD-**lal**, 58%; LDL-**vfc'**, 85%

(i) MsCl, DMAP, Et_3N
CH_2Cl_2, 0 ºC, 1 h

(ii) $^iPr_2NEt$, DMAP, pyridine
$C_2H_4Cl_2$, 25ºC, 12 h
(iii) BBr_3, CH_2Cl_2
-40 ºC, 5 min (for LLD-ffs')



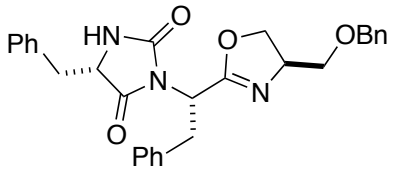
1

LLD-**1ffs'**, 45%; LLL-**1fff**, 39%
DLD-**1lal**, 41%; LDL-**1lal**, 40%
LLD-**1lal**, 39%; LDL-**1vfc'**, 49%
LLD-**1ffs**, 89%

A representative synthesis of LLD-**1ffs** is described as follows: Compound LLD-**9ffs'** (86 mg, 0.17 mmol) and DMAP (2 mg, 0.02 mmol) were dissolved in dichloromethane at 0 ºC under Ar. Then, Et_3N (0.07 mL, 0.52 mmol) and methanesulfonyl chloride (0.02 mL, 0.21 mmol) were added successively. Solution was stirred for 1 h at 0 ºC until starting material was disappeared as monitored by TLC. The organic layer was washed with water, brine and dried over $MgSO_4$, then was completely removed by vacuum. Crude oil was dissolved in 1,2-dichloroethane (2 mL), then DMAP (2 mg, 0.02 mmol), $^iPr_2NEt$ (0.09 mL, 0.52 mmol) and pyridine (0.04 mL, 0.52 mmol) were added together. Solution was stirred at room temperature for 12 h, and then was washed with 10 % citric acid, saturated $NaHCO_3$ and brine. Organic solution was dried over $MgSO_4$ and removed under reduce pressure. The final product was purified by silica gel column

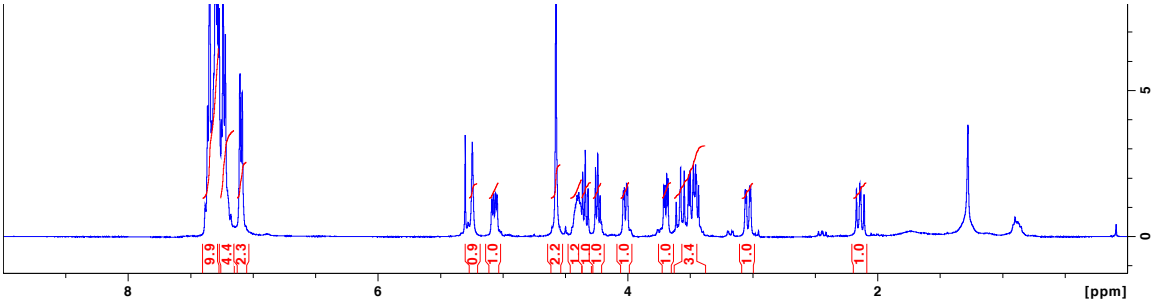
chromatography using hexanes:ethyl acetate (1.5:1) to afford 37 mg (45 %) of LLD-**1ffs'** as a white solid. Compound LLD-**1ffs'** (4.7 mg, 0.01 mmol) was dissolved in dichloromethane (1.3 mL) under Ar atmosphere. The solution was cooled down to -40 ºC, then the solution of BBr_3 (0.03 mL, 0.03 mmol) was added dropwise within 5 min and the reaction was stirred for 5 min. The reaction was quenched with saturated $NaHCO_3$. The organic phase was separated, and extracted with saturated $NaHCO_3$ and saturated NH_4Cl. The organic layer was collected, dried over $MgSO_4$ and removed under reduce pressure. Crude oil was purified by column chromatography using 5 % MeOH/CH_2Cl_2 to obtain product LLD-**1ffs** (3.4 mg, 89 %) as a solid.

(*S*)-5-Benzyl-3-((*S*)-1-((*R*)-4-((benzyloxy)methyl)-4,5-dihydrooxazol-2-yl)-2-phenylethyl)imidazolidine-2,4-dione (LLD-1ffs')

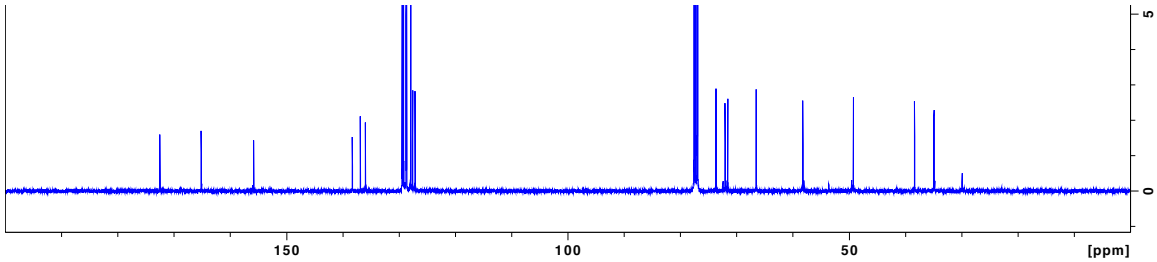


Data for LLD-1ffs'. White solid, 45 % yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.41-7.15 (m, 13H), 7.10 (m, 2H), 5.24 (br, 1H), 5.07 (dd, *J* = 11.6, 4.6 Hz, 1H), 4.58 (s, 2H), 4.46-4.37 (m, 1H), 4.34 (m, 1H), 4.24 (m, 1H), 4.02 (dd, *J* = 10.5, 3.0 Hz, 1H), 3.70 (dd, *J* = 9.5, 4.1 Hz, 1H), 3.61-3.43 (m, 3H), 3.03 (dd, *J* = 13.9, 3.5 Hz, 1H), 2.14 (dd, *J* = 13.8, 10.6 Hz, 1H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 172.5, 165.2, 155.8, 138.3, 136.9, 136.0, 129.4, 129.2, 128.7, 128.6, 127.9, 127.9, 127.6, 127.1, 73.7, 72.0, 71.5, 66.5, 58.2, 49.2, 38.4, 34.9; HRMS (ESI) m/z calcd for $C_{29}H_{30}N_3O_4^+$ 484.2231; found 484.2246 $(M+H)^+$

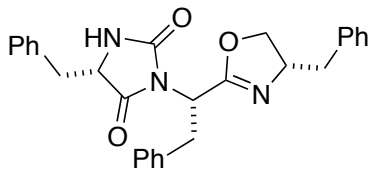
^{1}H NMR spectrum:



^{13}C NMR spectrum:

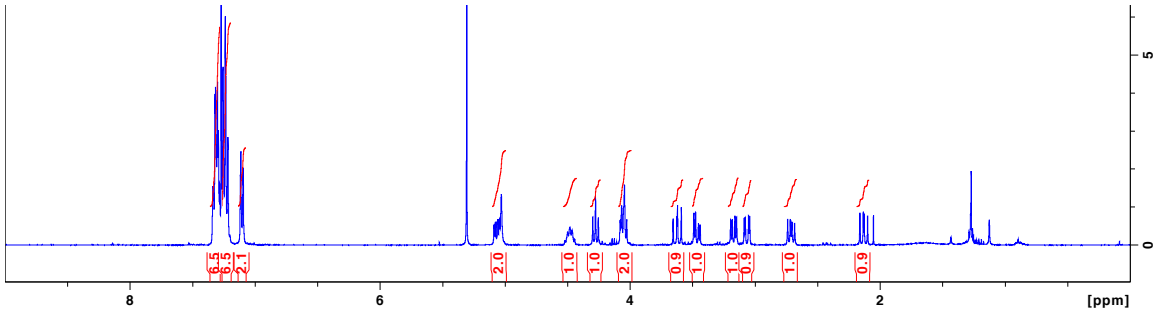


(*S*)-5-Benzyl-3-((*S*)-1-((*S*)-4-benzyl-4,5-dihydrooxazol-2-yl)-2-phenylethyl)imidazolidine-2,4-dione (LLL-1fff)

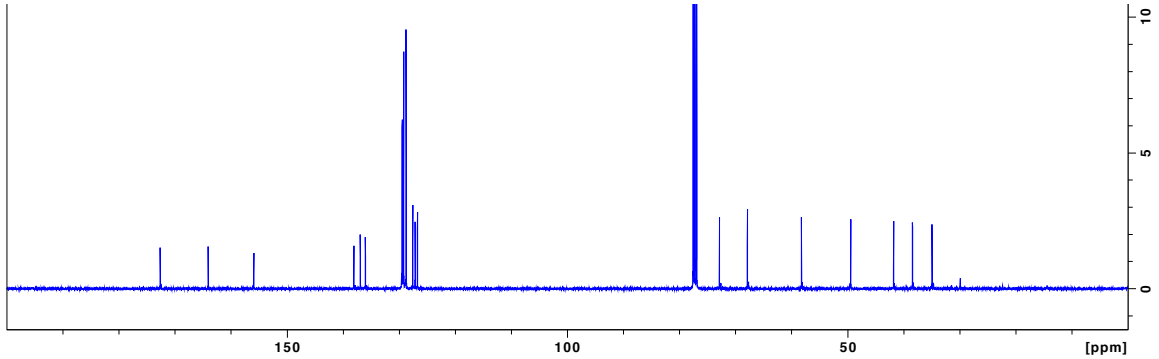


Data for LLL-1fff. Solid, 39 % yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.35-7.20 (m, 13H), 7.11 (m, 2H), 5.07 (m, 1H), 5.03 (br, 1H), 4.48 (m, 1H), 4.28 (t, *J* = 8.9 Hz, 1H), 4.1-4.01 (m, 2H), 3.62 (dd, *J* = 14.1, 11.8 Hz, 1H), 3.47 (dd, *J* = 14.2, 5.2 Hz, 1H), 3.17 (dd, *J* = 13.8, 5.3 Hz, 1H), 3.06 (dd, *J* = 13.8, 3.6 Hz, 1H), 2.71 (dd, *J* = 13.8, 8.6 Hz, 1H), 2.13 (dd, *J* = 13.9, 10.7 Hz, 1H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 172.6, 164.1, 155.9, 138.1, 136.9, 136.0, 129.5, 129.4, 129.2, 128.8, 127.6, 127.2, 126.7, 72.8, 67.9, 58.2, 49.4, 41.8, 38.4, 34.9; HRMS (ESI) m/z calcd for $C_{28}H_{28}N_3O_3^+$ 454.2125; found 454.2153 $(M+H)^+$

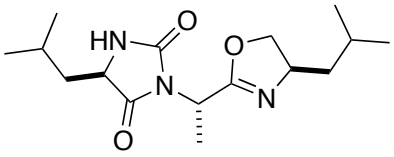
^{1}H NMR spectrum:



^{13}C NMR spectrum:

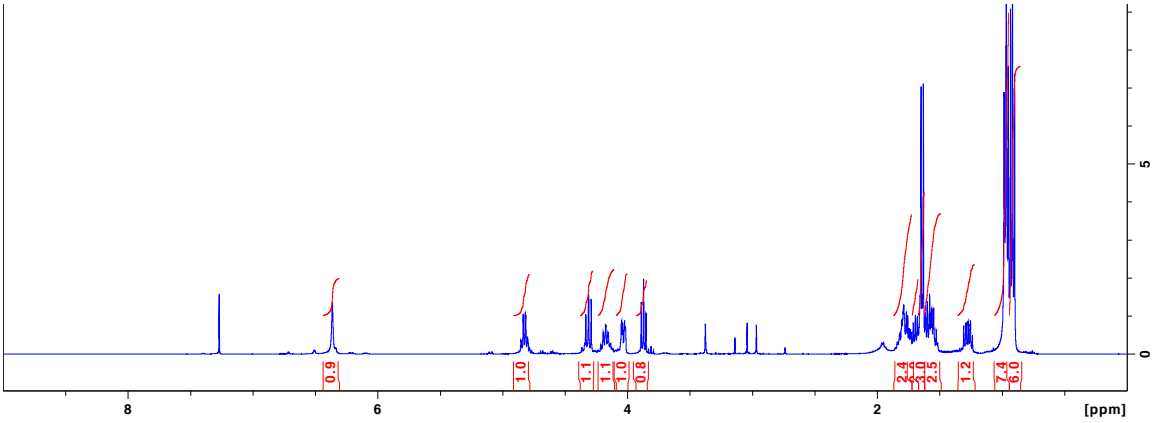


(*R*)-5-Isobutyl-3-((*S*)-1-((*R*)-4-isobutyl-4,5-dihydrooxazol-2-yl)ethyl)imidazolidine-2,4-dione (DLD-1lal)

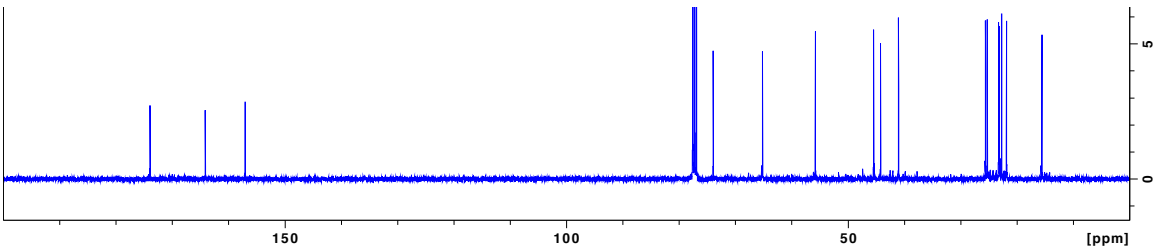


Data for DLD-1lal. Oil, 41 % yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 6.36 (br, 1H), 4.82 (m, 1H), 4.31 (m, 1H), 4.18 (m, 1H), 4.04 (m, 1H), 3.87 (t, *J* = 7.6 Hz, 1H), 1.87-1.67 (m, 3H), 1.64 (d, *J* = 7.2 Hz, 3H), 1.62-1.50 (m, 2H), 1.28 (m, 1H), 0.97 (m, 6H), 0.92 (m, 6H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 173.9, 164.1, 157.0, 73.9, 65.2, 55.8, 45.4, 44.2, 41.0, 25.6, 25.3, 23.3, 23.2, 22.7, 21.8, 15.5; HRMS (ESI) m/z calcd for $C_{16}H_{28}N_3O_3^+$ 310.2125; found 310.2142 $(M+H)^+$

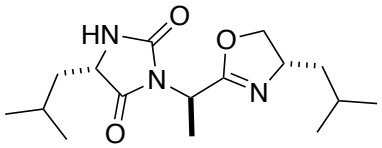
^{1}H NMR spectrum:



^{13}C NMR spectrum:

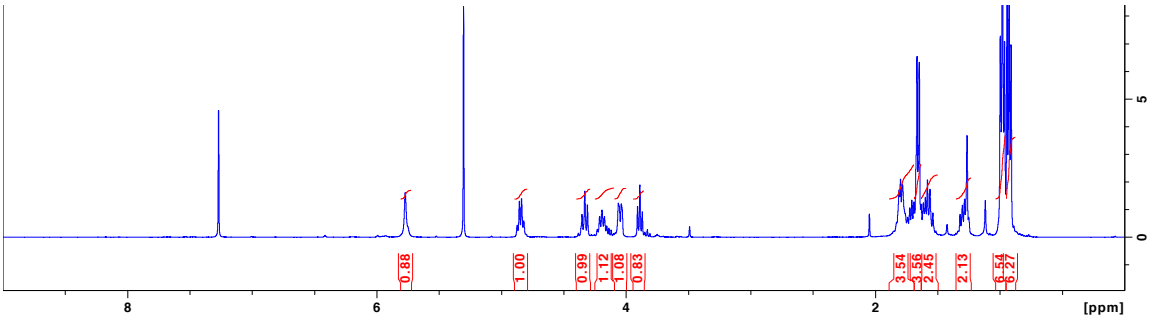


(*S*)-5-*Iso*butyl-3-((*R*)-1-((*S*)-4-*iso*butyl-4,5-dihydrooxazol-2-yl)ethyl)imidazolidine-2,4-dione (LDL-1lal)

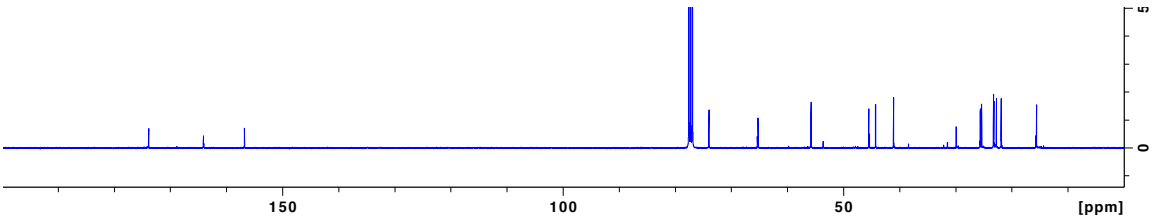


Data for LDL-1lal. Oil, 40 % yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 5.77 (s, 1H), 4.85 (q, *J* = 7.1 Hz, 1H) 4.33 (m, 1H), 4.25-4.09 (m, 1H), 4.05 (m, 1H), 3.89 (t, *J* = 7.6 Hz, 1H), 1.85-1.75 (m, 2H), 1.66 (d, *J* = 7.2 Hz, 3H), 1.62-1.53 (m, 2H), 1.35-1.24 (m, 2H), 0.98 (t, *J* = 6.8 Hz, 6H), 0.93 (t, *J* = 6.4 Hz, 6H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 173.8, 164.1, 156.8, 74.0, 65.2, 55.8, 45.5, 44.3, 41.1, 29.9, 25.6, 25.4, 23.3, 22.7, 21.9, 15.6; HRMS (ESI) m/z calcd for $C_{16}H_{28}N_3O_3^+$ 310.2125; found 310.2140 $(M+H)^+$

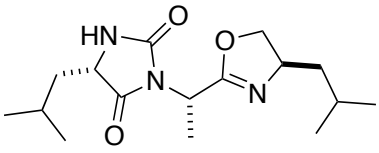
^{1}H NMR spectrum:



^{13}C NMR spectrum:

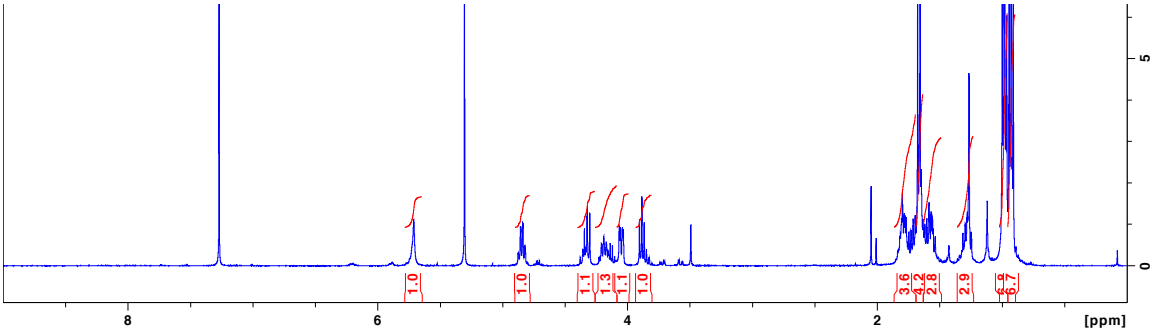


(*S*)-5-*Iso*butyl-3-((*S*)-1-((*R*)-4-*iso*butyl-4,5-dihydrooxazol-2-yl)ethyl)imidazolidine-2,4-dione (LLD-1lal)

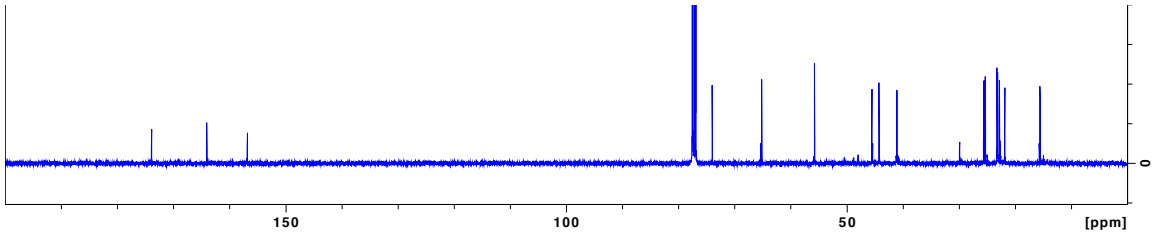


Data for LLD-1lal. Oil, 39 % yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 5.71 (br, 1H), 4.84 (m, 1H), 4.38-4.30 (m, 1H), 4.24-4.15 (m, 1H), 4.05 (m, 1H), 3.88 (m, 1H), 1.86-1.76 (m, 2H), 1.67 (d, *J* = 7.2 Hz, 3H), 1.61-1.52 (m, 2H), 1.36-1.22 (m, 2H), 0.98 (m, 6H), 0.93 (m, 6H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 173.9, 164.1, 156.8, 74.0, 65.2, 55.8, 45.5, 44.3, 41.1, 25.6, 25.3, 23.3, 23.1, 22.8, 21.8, 15.6; HRMS (ESI) m/z calcd for $C_{16}H_{28}N_3O_3^+$ 310.2125; found 310.2120 $(M+H)^+$

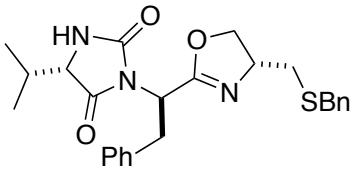
^{1}H NMR spectrum:



^{13}C NMR spectrum:

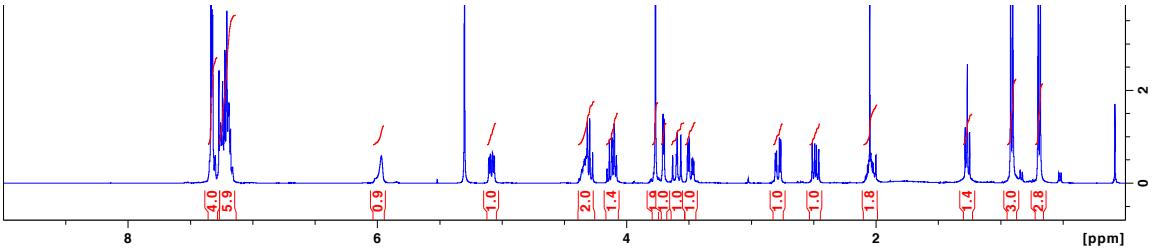


(*S*)-3-((*R*)-1-((*R*)-4-((Benzylthio)methyl)-4,5-dihydrooxazol-2-yl)-2-phenylethyl)-5-isopropylimidazolidine-2,4-dione (LDL-1vfc')

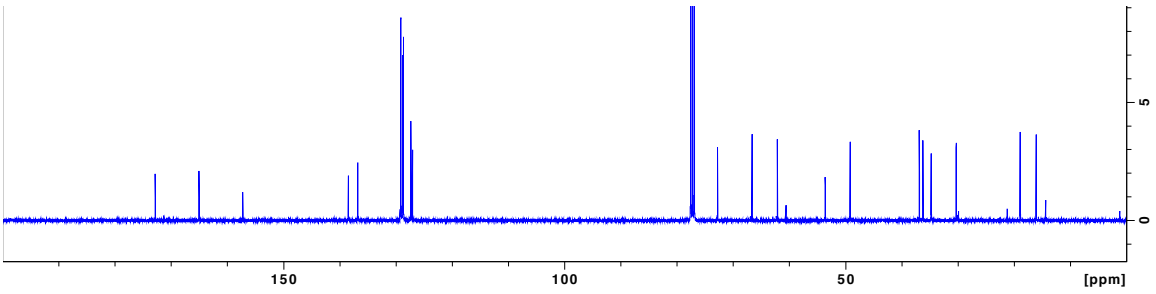


Data for LDL-1vfc'. Yellow solid, 49 % yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.33-7.13 (m, 10H), 5.97 (br, 1H), 5.09 (m, 1H), 4.39-4.26 (m, 2H), 4.16-4.08 (m, 1H), 3.77 (s, 2H), 3.71 (m, 1H), 3.60 (dd, *J* = 14.5, 11.7 Hz, 1H), 3.49 (dd, *J* = 14.5, 5.1 Hz, 1H), 2.79 (dd, *J* = 13.4, 4.4 Hz, 1H), 2.48 (dd, *J* = 13.4, 7.9 Hz, 1H), 2.10-2.00 (m, 1H), 0.91 (d, *J* = 7.0 Hz, 3H), 0.69 (d, *J* = 6.8 Hz, 3H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 172.8, 165.0, 157.3, 138.5, 136.8, 129.1, 129.0, 128.8, 128.7, 127.3, 127.0, 66.6, 62.1, 53.6, 49.2, 36.9, 36.2, 34.7, 30.3, 18.9, 16.1; HRMS (ESI) m/z calcd for $C_{25}H_{30}N_3O_3S^+$ 452.2002; found 452.2041 $(M+H)^+$

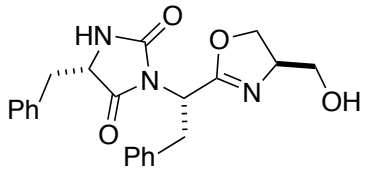
^{1}H NMR spectrum:



^{13}C NMR spectrum

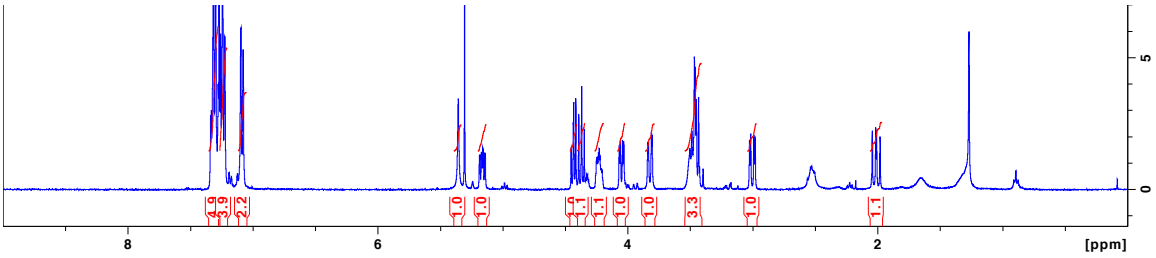


(*S*)-5-Benzyl-3-((*S*)-1-((*R*)-4-(hydroxymethyl)-4,5-dihydrooxazol-2-yl)-2-phenylethyl)imidazolidine-2,4-dione (LLD-1ffs)

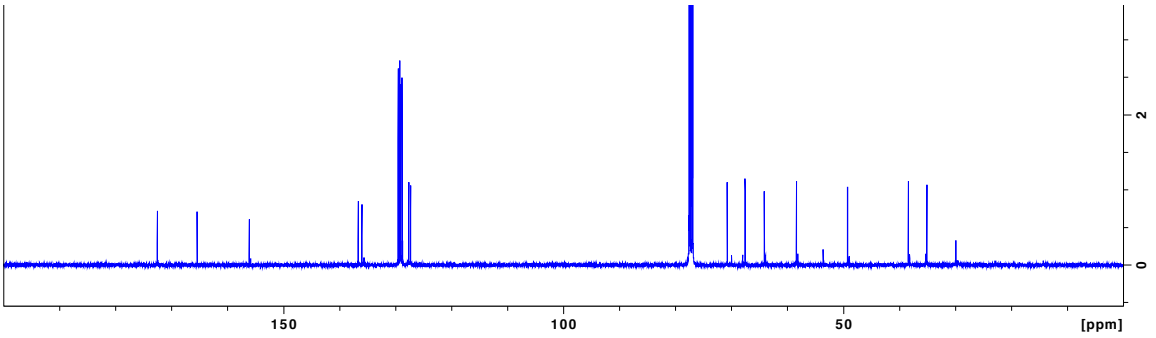


Data for LLD-1ffs. Solid, 89 % yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.36-7.22 (m, 8H), 7.09 (m, 2H), 5.36 (s, 1H), 5.16 (m, 1H), 4.44 (m, 1H), 4.40 (m, 1H), 4.27-4.20 (m, 1H), 4.05 (m, 1H), 3.82 (m, 1H), 3.55-3.41 (m, 3H), 3.00 (dd, *J* = 13.8, 3.7 Hz, 1H), 2.54 (m, 1H), 2.01 (dd, *J* = 13.8, 10.9 Hz, 1H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 172.5, 165.4, 156.1, 136.7, 136.0, 129.5, 129.2, 129.1, 128.8, 127.6, 127.3, 70.7, 67.6, 64.1, 58.4, 49.3, 38.4, 35.1; HRMS (ESI) m/z calcd for $C_{22}H_{24}N_3O_4^+$ 394.1716; found 394.1787 $(M+H)^+$

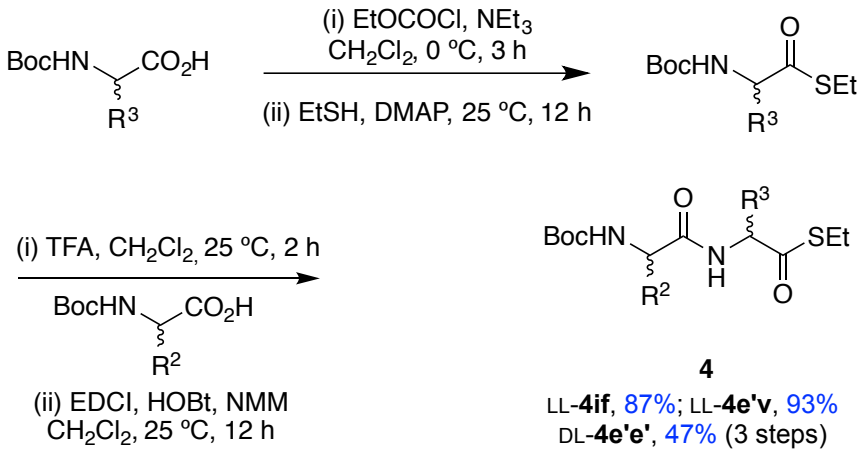
^{1}H NMR spectrum:



^{13}C NMR spectrum



*General Preparation of Ethyl-N-Boc-Dipeptide Thioester **4***



4

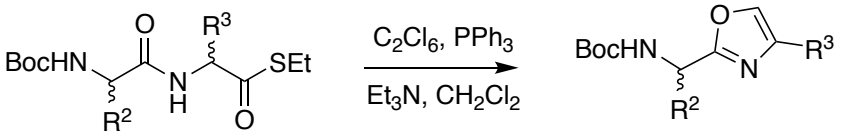
LL-**4if**, 87%; LL-**4e'v**, 93%
DL-**4e'e'**, 47% (3 steps)

A representative synthesis of compound LL-**4e'v** is described as follows: Boc-valine (1.52 g, 7 mmol) was dissolved in dichloromethane (40 mL) at 0 ºC under Ar. Et_3N (6.83 mL, 49 mmol) was added in one portion followed by ethyl chloroformate (2.40 mL, 25 mmol) over 15 min. DMAP (0.09 g, 0.7 mmol) and EtSH (2.59 mL, 35 mmol) were added subsequently and the reaction was stirred for additional 2 h at 0 ºC. The reaction was quenched using 2.0 mL of glacial acetic acid. The mixture was then evaporated under reduced pressure and redissolved in dichloromethane. Solution was washed 3 times with 10 % citric acid, 3 times with saturated $NaHCO_3$, brine and dried over $MgSO_4$. Solvent was removed to obtain crystalline solid, which was used for the next step without further purification. Boc deprotection of the Boc-valine-SEt (1.02 g, 3.89 mmol) was performed according to the Method B of the General Procedure of Boc Deprotection as described above.

Solution of Boc-L-glutamic acid γ-benzyl ester (1.31 g, 3.89 mmol) and HOBt (0.58 g, 4.28 mmol) were dissolved in dichloromethane (50 mL) at 0 ºC. EDCI (0.82 g, 4.28 mmol) was added and the solution was stirred for 30 min. *N*-methyl morpholine (1.28

mL, 11.7 mmol) and the NH_2-valine-SEt from the previous step dissolved in dichloromethane (20 mL) were added, and the reaction was stirred at ambient temperature for 12 h under Ar. Solution was washed with 10 % citric acid, saturated $NaHCO_3$, brine and dried over $MgSO_4$, then was removed under reduced pressure to obtain compound **4** (1.74 g, 93 %) as solid.

*General Preparation of N-Boc Oxazole **6***



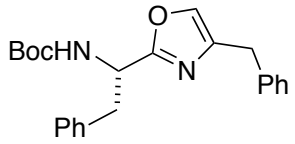
4 **6**

L-**6ff**, 40%; L-**6if**, 37%
L-**6e'v**, 57%; D-**6e'e'**, 39%

A representative synthesis of L-**6e'v** is described as follows: Dipeptide **4** (1.16 g, 2.41 mmol) was added with 10 % Pd/C (0.12 g, 0.02 mmol) and small amount of $MgSO_4$ in round-bottom flask and flushed with Ar. Acetone (5 mL) was added into, then Et_3SiH (1.15 mL, 7.23 mmol) was added dropwise over 10 min. The solution was stirred for 30 min, and monitored the reaction by TLC until starting material disappeared. Once the reaction completed, the solution was filtered through celite and evaporated under reduced pressure to obtain crude aldehyde **5** as oil. In the round-bottom flask, hexachloroethane (1.71 g, 7.23 mmol) and PPh_3 (1.90 g, 7.23 mmol) were dissolved in acetonitrile (19 mL) in an ice bath. Aldehyde **5** in 7 mL acetonitrile was added dropwise and the solution was stirred for 10 min at 0 ºC. Et_3N (2.02 mL, 14.5 mmol) was added to the solution, and the reaction was stirred at ambient temperature for 12 h. Solution was removed under reduced pressure, and then redissolved in ethyl acetate. Organic solution was washed with water 3 times, brine and dried over $MgSO_4$, and then dried

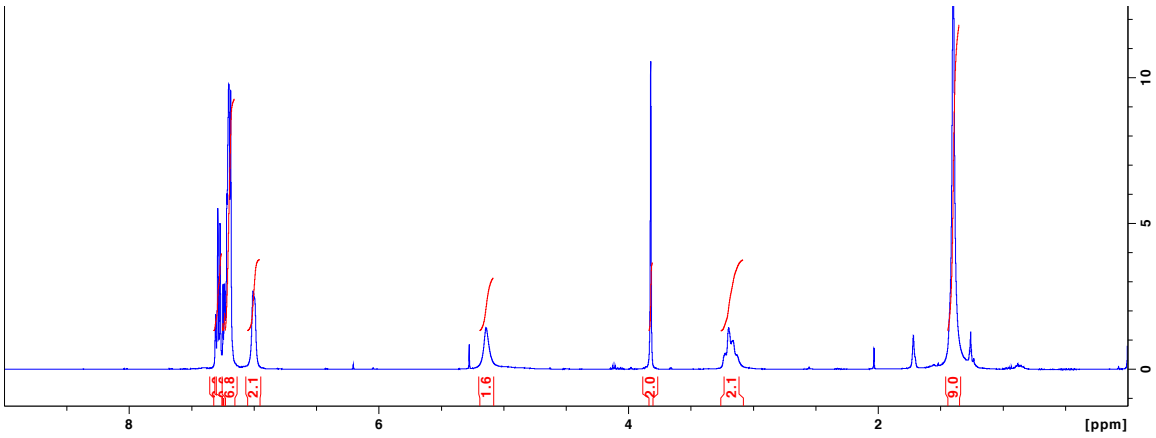
completely. Crude was purified by silica gel column chromatography using hexanes:ethyl acetate to obtain oxazole **6** (0.49 g, 57 %) as solid.

***tert*-Butyl (*S*)-(1-(4-benzyloxazol-2-yl)-2-phenylethyl)carbamate (L-6ff)**

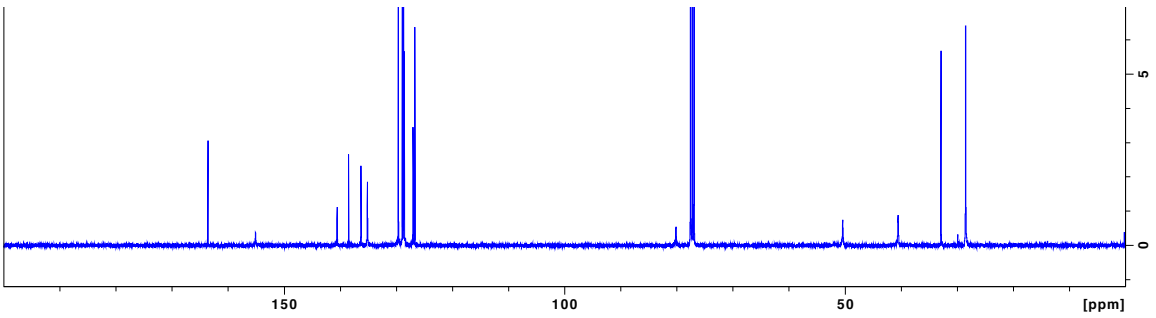


Data for L-6ff. Solid, 40 % yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.32-7.17 (m, 10H), 7.01 (m, 2H), 5.14 (br, 1H), 3.82 (s, 2H), 3.25-3.11 (m, 2H), 1.40 (s, 9H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 163.6, 155.1, 140.5, 138.5, 136.3, 135.1, 129.6, 128.9, 128.7, 128.6, 127.0, 126.7, 80.1, 50.4, 40.5, 32.9, 28.5; HRMS (ESI) m/z calcd for $C_{23}H_{27}N_2O_3^+$ 379.2016; found 379.1984 $(M+H)^+$

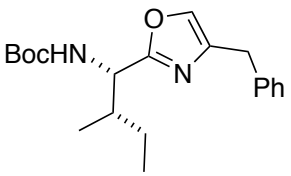
^{1}H NMR spectrum:



^{13}C NMR spectrum:

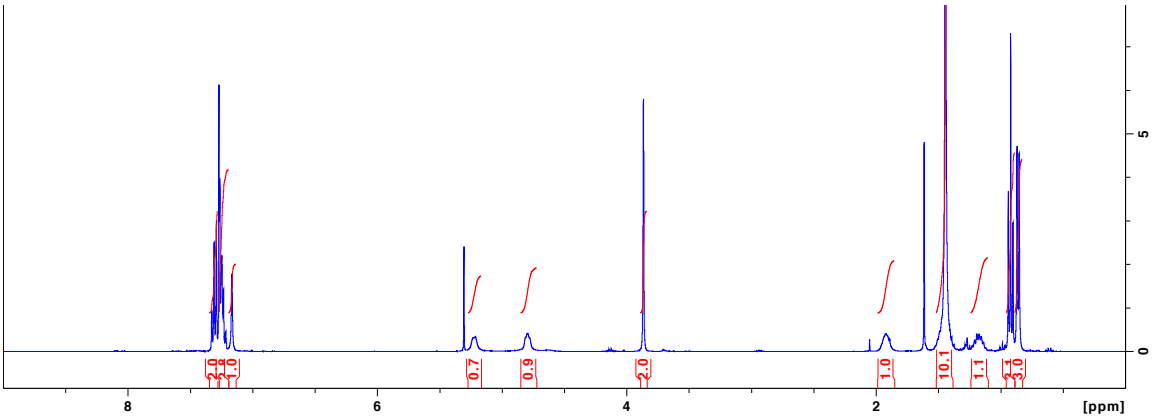


***tert*-Butyl ((1*S*,2*S*)-1-(4-benzyloxazol-2-yl)-2-methylbutyl)carbamate (L-6if)**

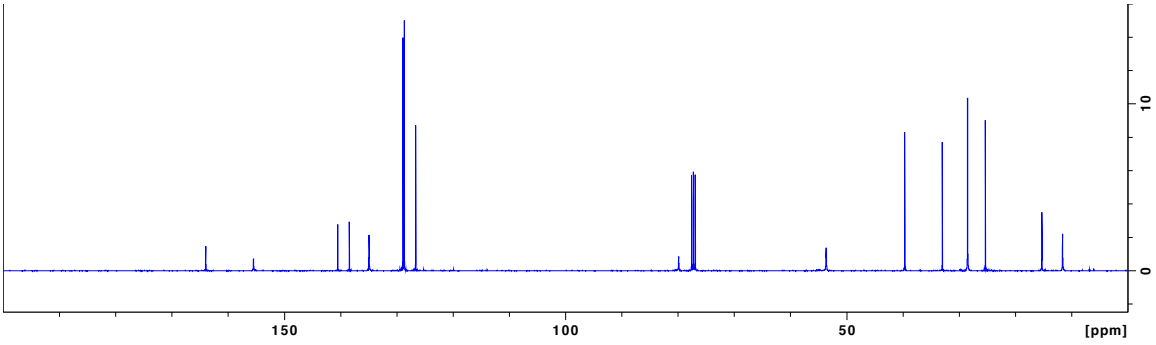


Data for L-6if. Solid, 37 % yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.33-7.21 (m, 5H), 7.17 (s, 1H), 5.22 (m, 1H), 4.80 (m, 1H), 3.87 (s, 2H), 1.90 (m, 1H), 1.52-1.38 (m, 1H), 1.44 (s, 9H), 1.18 (m, 1H), 0.92 (t, *J* = 7.4 Hz, 3H), 0.86 (d, *J* = 6.8 Hz, 3H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 164.0, 155.5, 140.5, 138.4, 134.9, 128.9, 128.6, 126.6, 79.9, 56.6, 39.7, 33.0, 28.5, 25.3, 15.3, 11.6; MS (ESI) m/z calcd for $C_{20}H_{29}N_2O_3^+$ 345.22; found 345.20 $(M+H)^+$

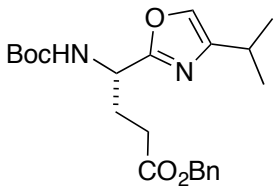
^{1}H NMR spectrum:



^{13}C NMR spectrum

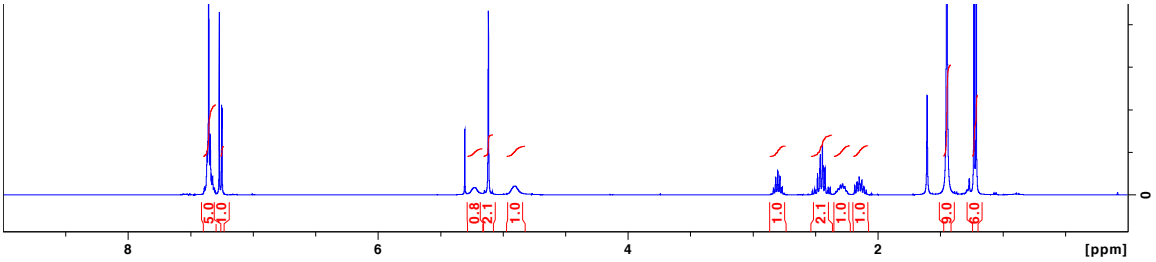


Benzyl (*S*)-4-((*tert*-butoxycarbonyl)amino)-4-(4-isopropyloxazol-2-yl)butanoate (L-6e'v)

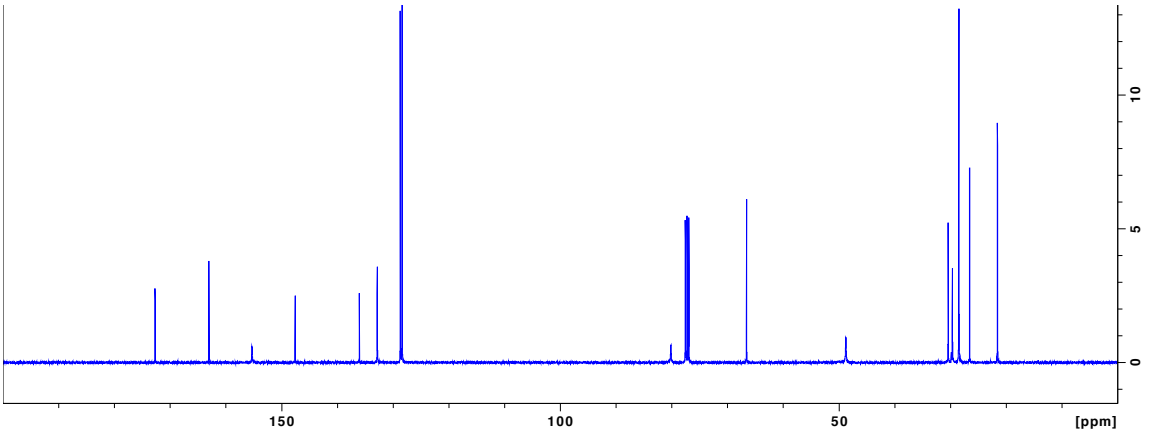


Data for L-6e'v Oil, 57 % yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.40-7.30 (m, 5H), 7.25 (d, *J* = 1.2 Hz, 1H), 5.23 (br, 1H), 5.12 (s, 2H), 4.91 (br, 1H), 2.80 (m, 1H), 2.54-2.37 (m, 2H), 2.35-2.22 (m 1H), 2.20-2.07 (m, 1H), 1.45 (s, 9H), 1.22 (d, *J* = 6.9 Hz, 6H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 172.7, 163.0, 155.3, 147.5, 136.0, 132.8, 128.7, 128.4, 80.1, 66.6, 48.7, 30.4, 29.6, 28.5, 26.5, 21.6, 21.5; HRMS (ESI) m/z calcd for $C_{22}H_{31}N_2O_5^+$ 403.2227; found 403.2164 $(M+H)^+$

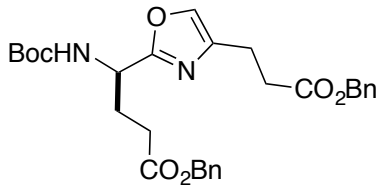
^{1}H NMR spectrum:



^{13}C NMR spectrum:

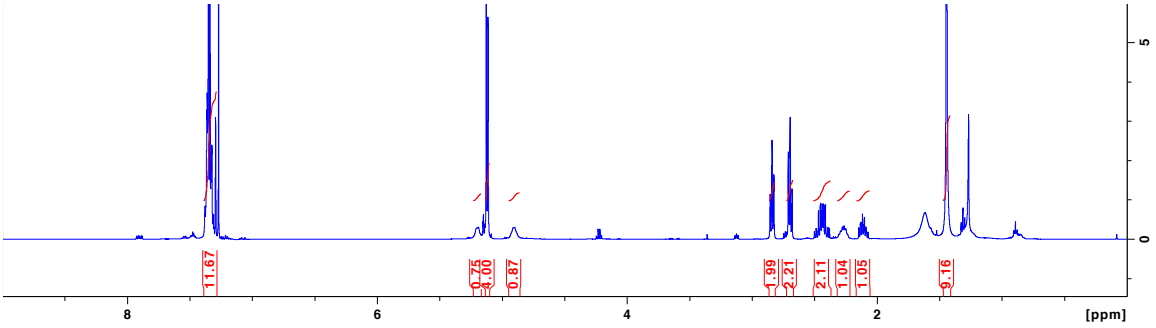


Benzyl (*R*)-4-(4-(3-(benzyloxy)-3-oxopropyl)oxazol-2-yl)-4-((*tert*-butoxycarbonyl)amino)butanoate (D-6e'e')

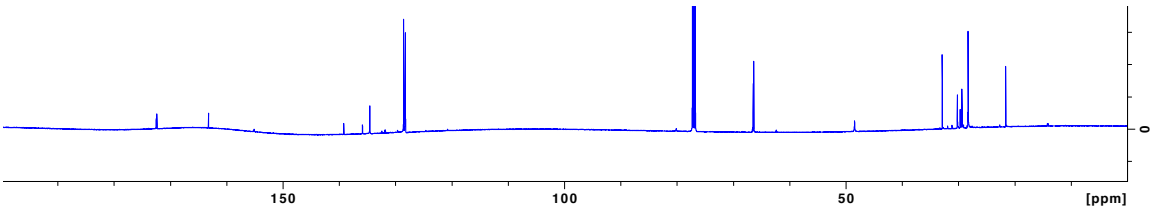


Data for D-6e'e' Oil, 39 % yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.39-7.28 (m, 11H), 5.20 (br, 1H), 5.12 (m, 4H), 4.91 (br, 1H), 2.84 (m, 2H), 2.70 (m, 2H), 2.51-2.37 (m, 2H), 2.32-2.22 (m, 1H), 2.17-2.05 (m, 1H), 1.44 (s, 9H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 172.5, 172.4, 163.2, 155.2, 139.2, 135.9, 135.8, 134.6, 128.6, 128.3, 128.2, 128.2, 80.2, 66.5, 66.4, 48.5, 32.9, 30.2, 29.7, 29.4, 28.3, 21.6; MS (ESI) m/z calcd for $C_{29}H_{35}N_2O_7^+$ 523.24; found 523.26 $(M+H)^+$

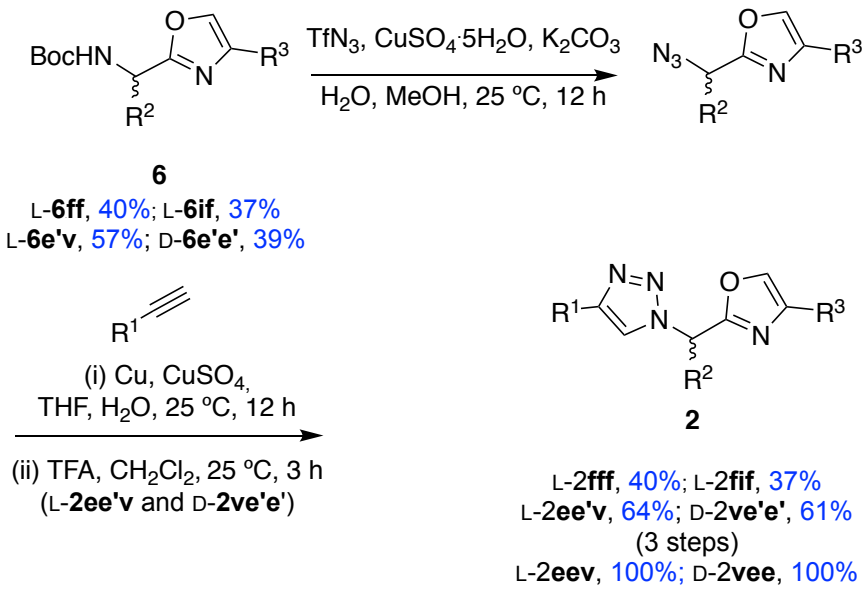
^{1}H NMR spectrum:



^{13}C NMR spectrum:



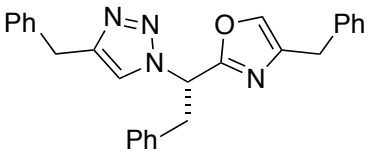
*General Procedure for the Preparation of Scaffold **2***



A representative synthesis of L-**2ee'v** is described as follows: Boc-deprotection of L-**6e'v** (218 mg, 0.54 mmol) was performed same as Method B of the General Procedure of Boc Deprotection as described above. Crude oxazole was dissolved in methanol (3.5 mL), then $CuSO_4.5H_2O$ (1.25 mg, 0.005 mmol) and K_2CO_3 (112 mg, 0.81 mmol) dissolved in water (1.75 mL) was added into the reaction. Then, fresh TfN_3 (3 mmol) in dichloromethane was added directly to the reaction. Solution was stirred for 12 h at room temperature. Organic solvent was removed under vacuum at room temperature. The remaining liquid was diluted with water and pH was adjusted to 2 with concentrated HCl. The aqueous layer was extracted 4 times with ethyl acetate, then combined organic layer was washed with brine and dried over $MgSO_4$. The solvent was removed under reduced pressure to afford azido-oxazole, which was used in the next step without further purification. Azido-oxazole and 4-pentynoic acid (55.9 mg, 0.57 mmol) were dissolved in THF (1.7 mL) and water (0.3 mL). Subsequently, copper powder (34 mg, 0.54 mmol) and 1 M solution of $CuSO_4$ (0.05 mL, 0.05 mmol) were added. The solution

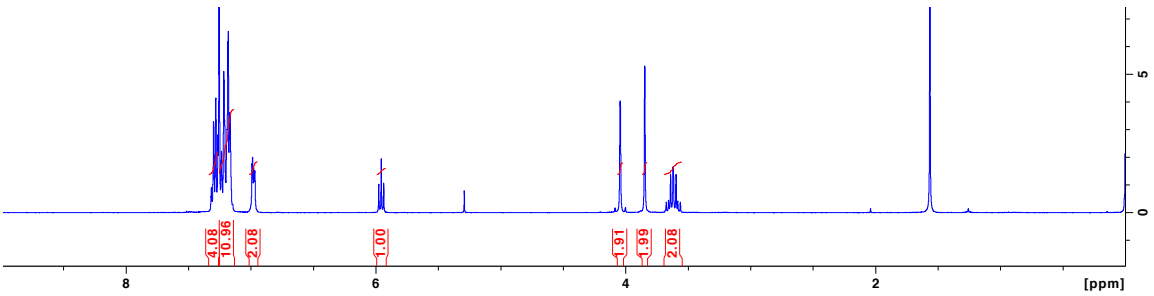
was stirred for 12 h. After the reaction was complete, pH was adjusted to 2 using concentrated HCl, then aqueous layer was extracted 4 times with ethyl acetate. Organic layer was combined, dried over $MgSO_4$ and removed under reduced pressure. Crude product was purified by silica gel column chromatography using hexanes : ethyl acetate (1:2) to obtain product L-**2ee'v** (148 mg, 64%) as solid. The hydrogenolysis of L-**2ee'v** and D-**2ve'e'** was performed as same as Method B in the General Procedure for the Preparation of scaffold **9** to obtain L-**2eev** and D-**2vee** as oil.

(*S*)-4-Benzyl-2-(1-(4-benzyl-1*H*-1,2,3-triazol-1-yl)-2-phenylethyl)oxazole (L-2fff)

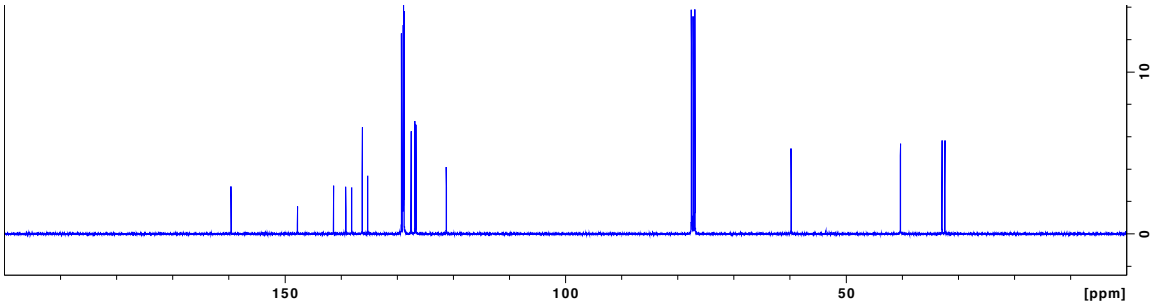


Data for L-2fff. Solid, 40 % yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.33-7.15 (m, 15H), 6.99 (m, 2H), 5.96 (t, *J* = 7.8 Hz, 1H), 4.04 (s, 2H), 3.85 (s, 2H), 3.62 (m, 2H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 159.6, 147.7, 141.3, 139.1, 138.1, 136.2, 135.2, 129.2, 128.9, 128.8, 128.8, 128.8, 128.7, 127.5, 126.8, 126.6, 121.2, 59.8, 40.3, 32.9, 32.4; HRMS (ESI) m/z calcd for $C_{27}H_{25}N_4O^+$ 421.2023; found 421.1885 $(M+H)^+$

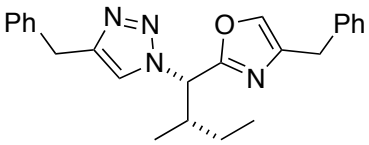
^{1}H NMR spectrum:



^{13}C NMR spectrum:

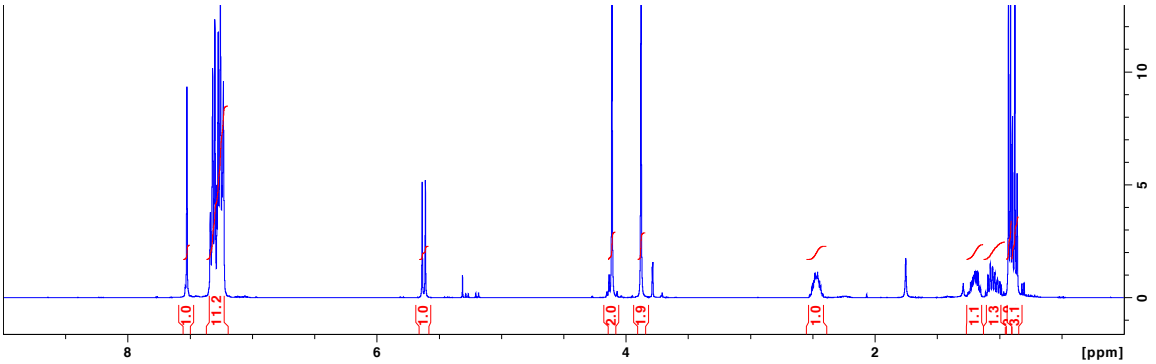


4-Benzyl-2-((1*S*,2*S*)-1-(4-benzyl-1*H*-1,2,3-triazol-1-yl)-2-methylbutyl)oxazole (L-2fif)

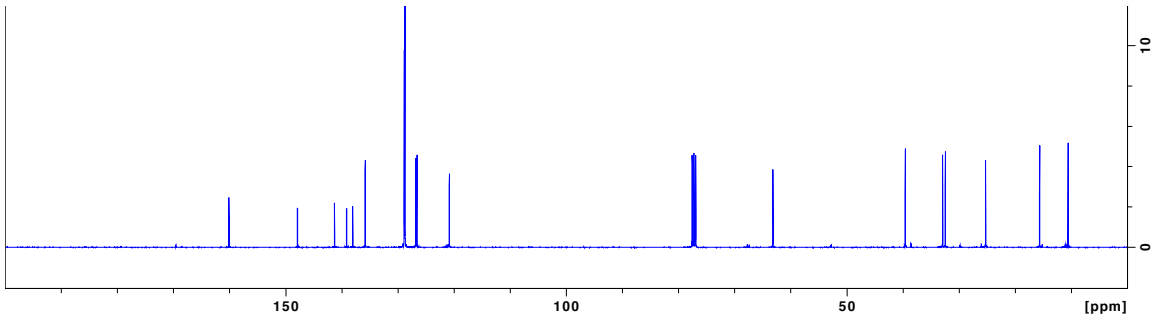


Data for L-2fif. Solid, 37 % yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.53 (s, 1H), 7.35-7.22 (m, 11H), 5.62 (d, *J* = 10.2 Hz, 1H), 4.11 (s, 2H), 3.88 (s, 2H), 2.53-2.41 (m, 1H), 1.26-1.14 (m, 1H), 1.12-0.97 (m, 1H), 0.92 (d, *J* = 6.7 Hz, 3H), 0.88 (t, *J* = 7.4 Hz, 3H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 160.1, 147.9, 141.3, 139.1, 138.0, 135.8, 128.9, 128.8, 128.7, 126.8, 126.6, 120.8, 63.2, 39.6, 32.9, 32.4, 25.3, 15.6, 10.5; HRMS (ESI) m/z calcd for $C_{24}H_{27}N_4O^+$ 387.2179; found 387.2030 $(M+H)^+$

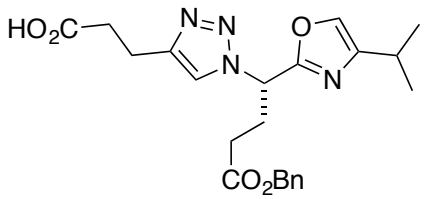
^{1}H NMR spectrum:



^{13}C NMR spectrum:

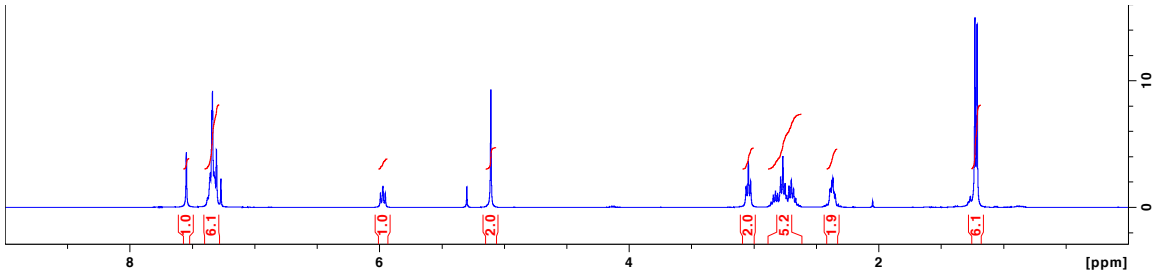


(*S*)-3-(1-(4-(Benzyloxy)-1-(4-isopropyloxazol-2-yl)-4-oxobutyl)-1*H*-1,2,3-triazol-4-yl)propanoic acid (L-2ee'v)

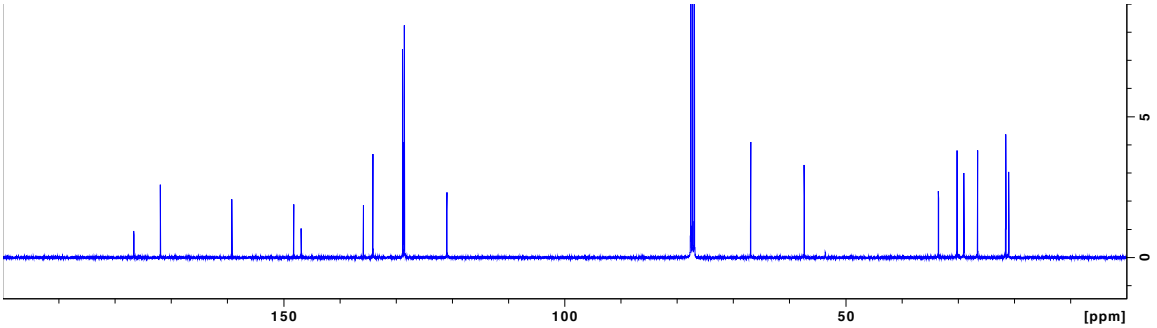


Data for L-2ee'v White solid, 64 % yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.55 (s, 1H), 7.39-7.29 (m, 6H), 5.967 (t, *J* = 7.8 Hz, 1H), 5.11 (s, 2H), 3.04 (t, *J* = 7.2 Hz, 2H), 2.87-2.66 (m, 5H), 2.37 (m, 2H), 1.22 (d, *J* = 6.8 Hz, 6H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 176.6, 171.9, 159.2, 148.2, 146.9, 135.8, 134.1, 128.8, 128.6, 128.5, 121.0, 66.9, 57.4, 33.5, 30.2, 29.0, 26.5, 21.5, 21.5, 21.0; HRMS (ESI) m/z calcd for $C_{22}H_{26}N_4O_5Na^+$ 449.1795; found 449.1795 $(M+Na)^+$

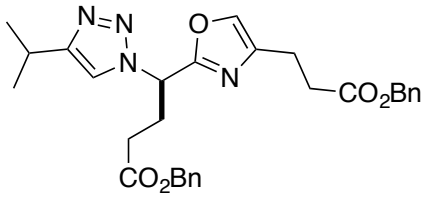
^{1}H NMR spectrum:



^{13}C NMR spectrum

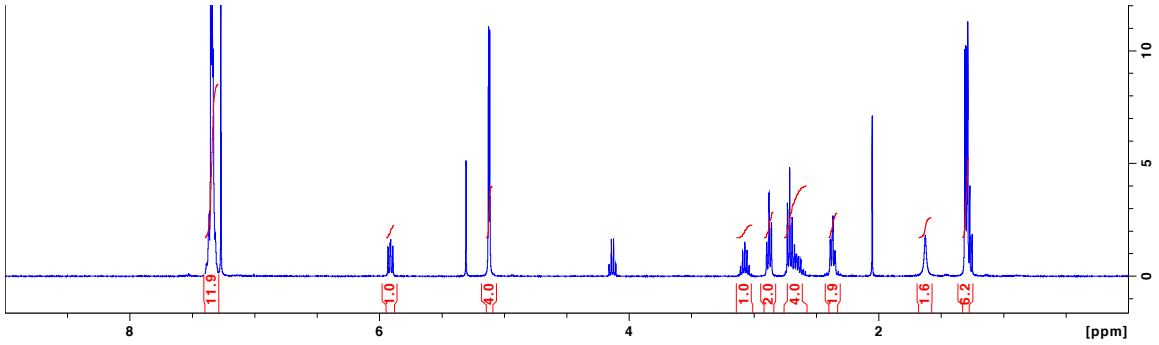


Benzyl (*R*)-4-(4-(3-(benzyloxy)-3-oxopropyl)oxazol-2-yl)-4-(4-isopropyl-1*H*-1,2,3-triazol-1-yl)butanoate (D-2ve'e')

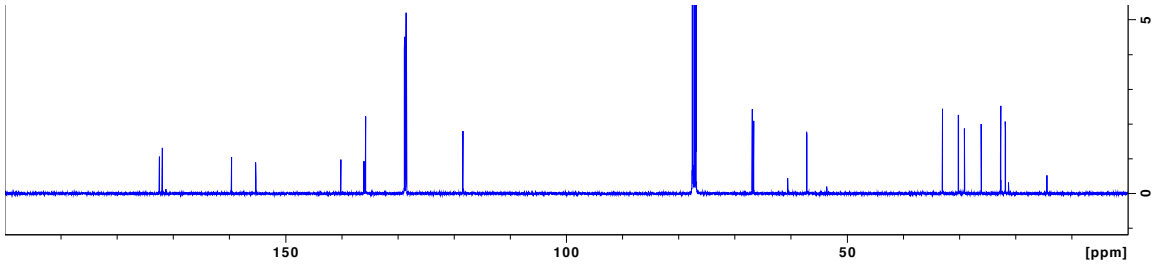


Data for D-2ve'e' White solid, 61 % yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.40-7.30 (m, 12H), 5.91 (dd, *J* = 8.8, 6.8 Hz, 1H), 5.13 (s, 2H), 5.12 (s, 2H), 3.07 (m, 1H), 2.88 (t, *J* = 7.3 Hz, 2H), 2.71 (t, *J* = 7.6 Hz, 2H**),** 2.68-2.58 (m, 2H), 2.40-2.33 (m, 2H), 1.31 (d, *J* = 2.9 Hz, 3H), 1.29 (d, *J* = 2.9 Hz, 3H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 172.5, 171.9, 159.6, 155.3, 140.1, 136.1, 135.8, 135.7, 128.8, 128.8, 128.6, 128.5, 128.4, 118.4, 66.9, 66.6, 57.2, 33.0, 30.2, 29.1, 26.1, 22.6, 22.6, 21.8; HRMS (ESI) m/z calcd for $C_{29}H_{33}N_4O_5^+$ 517.2445; found 517.2419 $(M+H)^+$

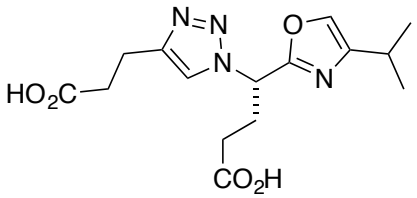
^{1}H NMR spectrum:



^{13}C NMR spectrum

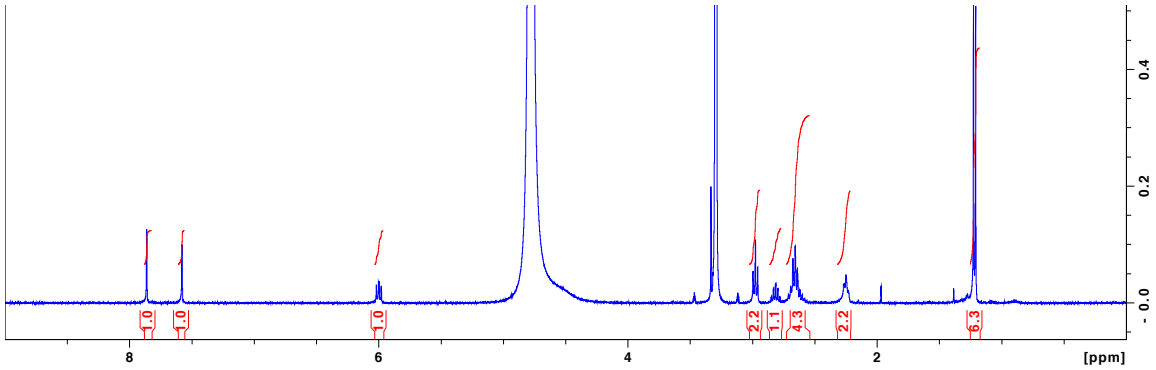


(*S*)-4-(4-(2-Carboxyethyl)-1*H*-1,2,3-triazol-1-yl)-4-(4-isopropyloxazol-2-yl)butanoic acid (L-2eev)

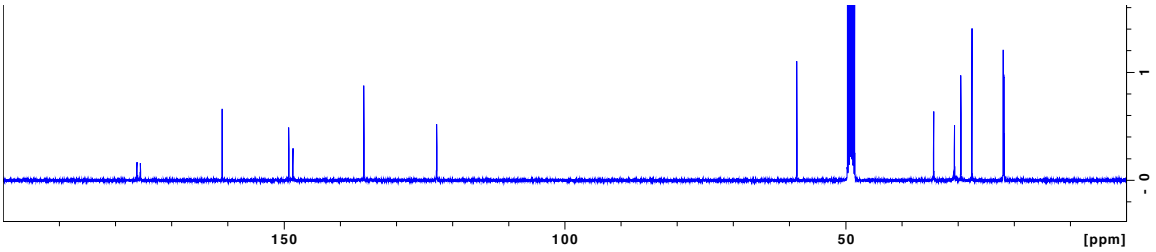


Data for L-2eev Colorless oil, 100 % yield. ^{1}H NMR (400 MHz, MeOD) δ 7.86 (s, 1H), 7.58 (s, 1H), 6.00 (dd, *J* = 8.9, 6.7 Hz, 1H), 2.98 (t, *J* = 7.4 Hz, 2H), 2.81 (m, 1H), 2.73-2.56 (m, 4H), 2.26 (m, 2H), 1.22 (d, *J* = 6.92 Hz, 6H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 176.2, 175.6, 161.0, 149.1, 148.4, 135.8, 122.8, 58.7, 34.3, 30.6, 29.5, 27.5, 21.9, 21.8, 21.7; HRMS (ESI) m/z calcd for $C_{15}H_{20}N_4O_5Na^+$ 359.1326; found 359.1319 $(M+Na)^+$

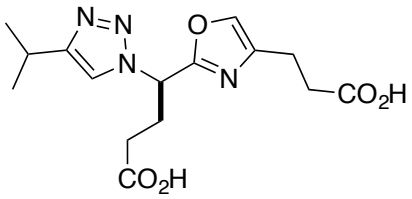
^{1}H NMR spectrum:



^{13}C NMR spectrum

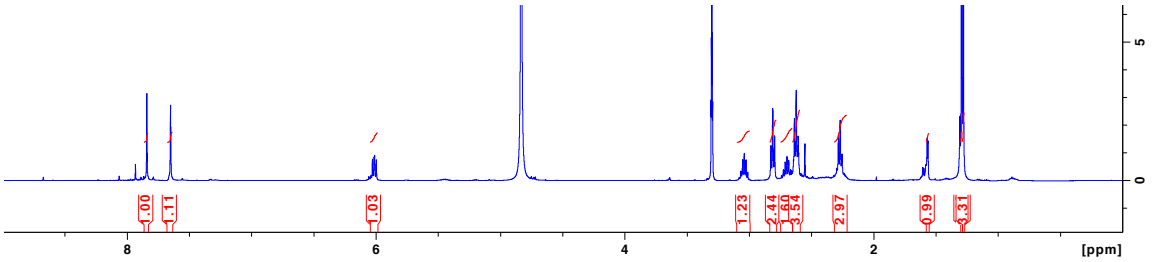


(*R*)-4-(4-(2-Carboxyethyl)oxazol-2-yl)-4-(4-isopropyl-1*H*-1,2,3-triazol-1-yl)butanoic acid (D-2vee)

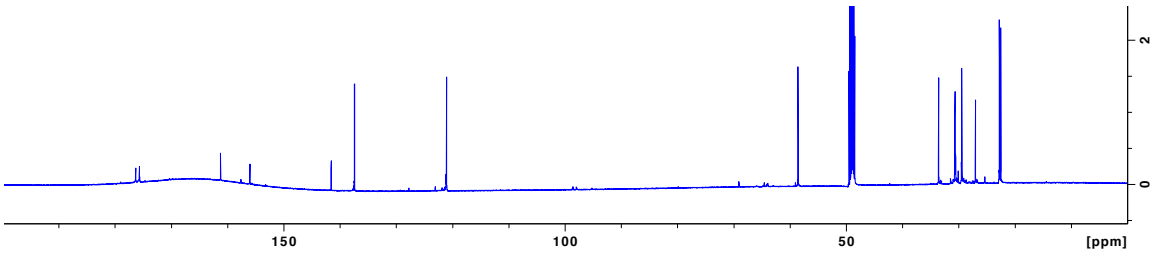


Data for D-2vee Colorless oil, 100 % yield. ^{1}H NMR (400 MHz, MeOD) δ 7.84 (s, 1H), 7.65 (s, 1H), 6.01 (dd, *J* = 9.0, 6.6 Hz, 1H), 3.04 (m, 1H), 2.81 (t, *J* = 7.4 Hz, 2H), 2.75-2.65 (m, 2H), 2.65-2.59 (m, 3H), 2.32-2.22 (m, 3H), 1.29 (m, 3H), 1.28 (m, 3H); ^{13}C NMR (100 MHz, MeOD) δ 176.3, 175.6, 161.2, 156.0, 141.6, 137.4, 121.0, 58.6, 33.6, 30.7, 29.5, 27.0, 22.8, 22.7, 22.5; HRMS (ESI) m/z calcd for $C_{15}H_{21}N_4O_5^+$ 337.1506; found 337.1532 $(M+H)^+$

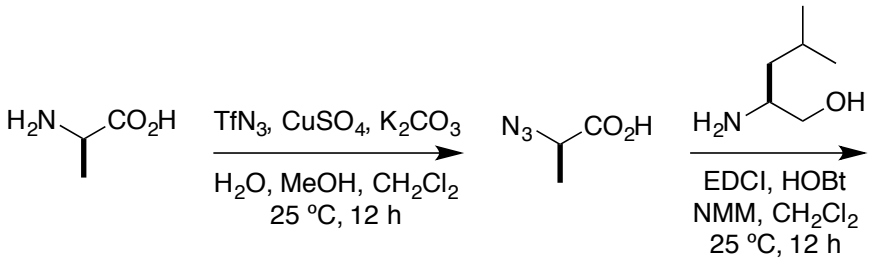
^{1}H NMR spectrum:

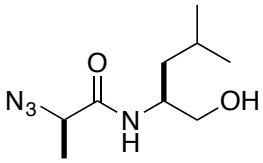


^{13}C NMR spectrum:



*Preparation of compound DL-**10al***





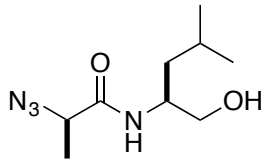
10

DL-**10al**, 39% (2 steps)

Alanine (0.45 g, 5.0 mmol) was dissolved in 30 mL MeOH, then the solution of $CuSO_4$ and K_2CO_3 in 15 mL water was added. Fresh solution of TfN_3 was added, and the reaction was stirred at room temperature for 12 h. Organic solvent was removed by rotavap, and then the aqueous phase was acidified with concentrated HCl until pH ~2.

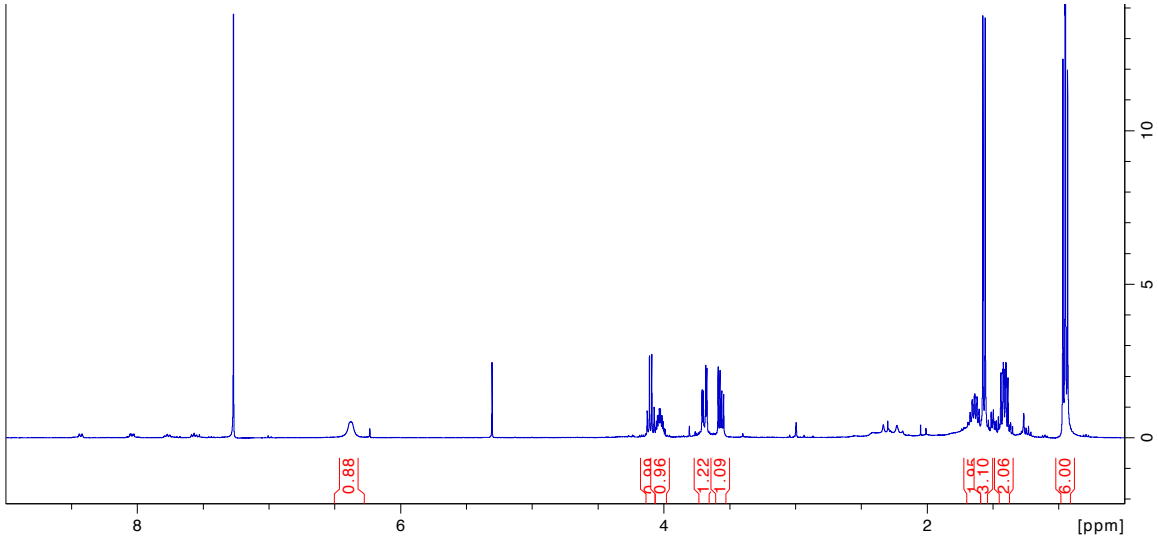
Solution was extracted 4 times with EtOAc, dried with $MgSO_4$, and evaporated to get the azido-Ala, which was used in the next step without further purification. To the solution of azido-Ala (5.0 mmol), HOBt (0.74 g, 5.5 mmol) and EDCI (1.05 g, 5.5 mmol) in dichloromethane (50 mL) was added with *N*-methyl morpholine (1.65 mL, 15 mmoL) and leucinol (0.59 g, 5.0 mmol) at 0 ºC. Reaction was stirred under Ar atmosphere for 12 h while allowing temperature to rise to room temperature. Organic layer was washed with 1N HCl (x3), Sat. $NaHCO_3$ (x3), and brine. Solvent was dried over $MgSO_4$, and removed under reduce pressure to obtain DL-**10al** as solid (0.42 g, 39%), which was reasonable pure enough without further purification.

(*R*)-2-Azido-*N*-((*S*)-1-hydroxy-4-methylpentan-2-yl)propanamide (DL-10al)

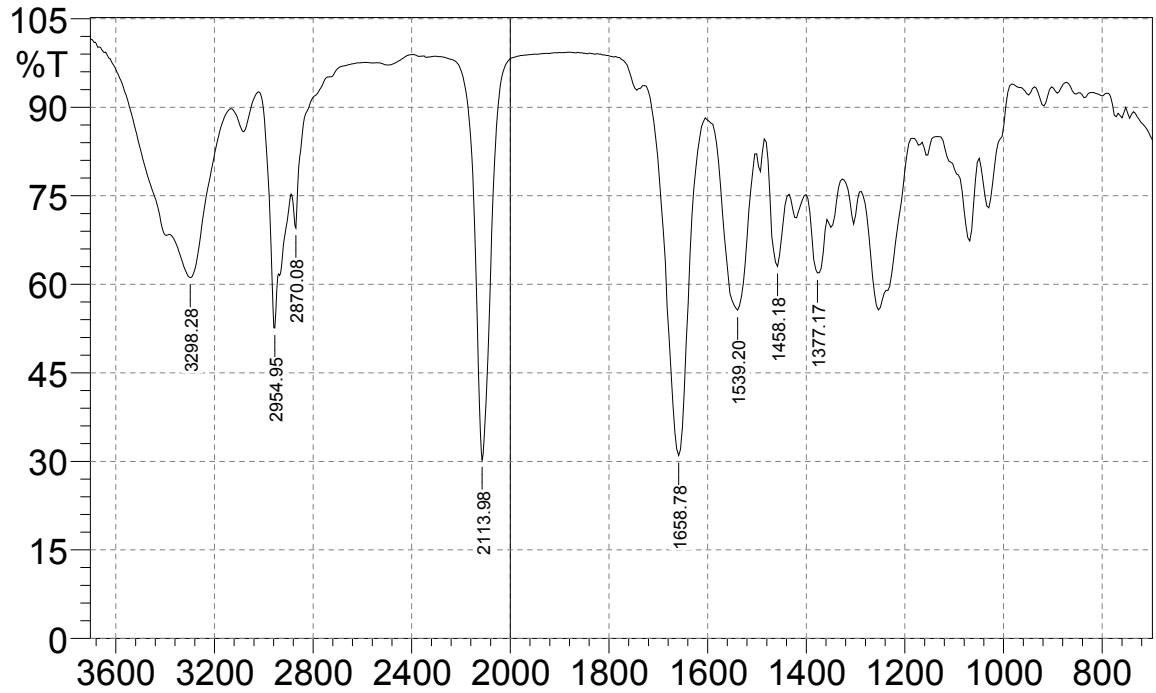


Data for DL-10al 39 % yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 6.38 (br, 1H), 4.10 (q, *J* = 7.0 Hz, 1H), 4.03 (m, 1H), 3.69 (dd, *J* = 11.0, 3.6 Hz, 1H), 3.57 (dd, *J* = 11.0, 5.7 Hz, 1H), 1.64 (m, 2H), 1.57 (d, *J* = 7.0 Hz, 3H), 1.42 (m, 2H), 0.96 (d, *J* = 6.7 Hz, 3H), 0.94 (d, *J* = 6.6 Hz, 3H)

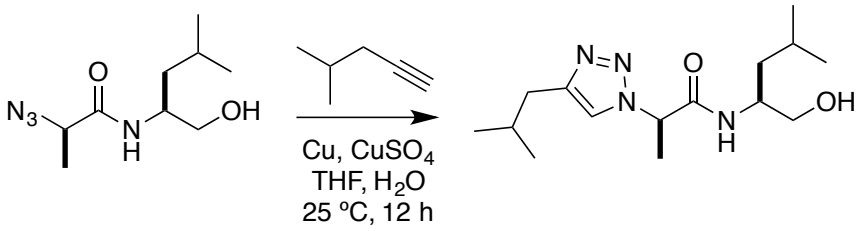
^{1}H NMR spectrum:



IR



Preparation of Compound DL-***11lal***

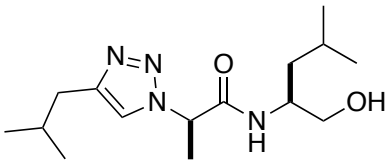


10 **11**

DL-**11lal**, 62%

Compound DL-**10al** (200 mg, 0.94 mmol) and 4-methylpentyne (0.12 mL, 0.98 mmol) were dissolved in 0.5 mL of water and 3 mL of THF. After that, the solution of 1M $CuSO_4$ (0.09 mL, 0.09 mmol) and Cu powder (60 mg, 0.94 mmol) was added, and the reaction was stirred for 12 h. Solvent was then completely removed by rotavap, and the crude was dissolved in EtOAc. Organic solvent was washed with 50 % NH_4OH (20 mL) and brine. Solution was dried over $MgSO_4$, evaporated, and the crude material was purified by flash column chromatography with EtOAc:hexanes (1.5:1 to 2.5:1 ratio) to obtain product as solid (173 mg, 62 %).

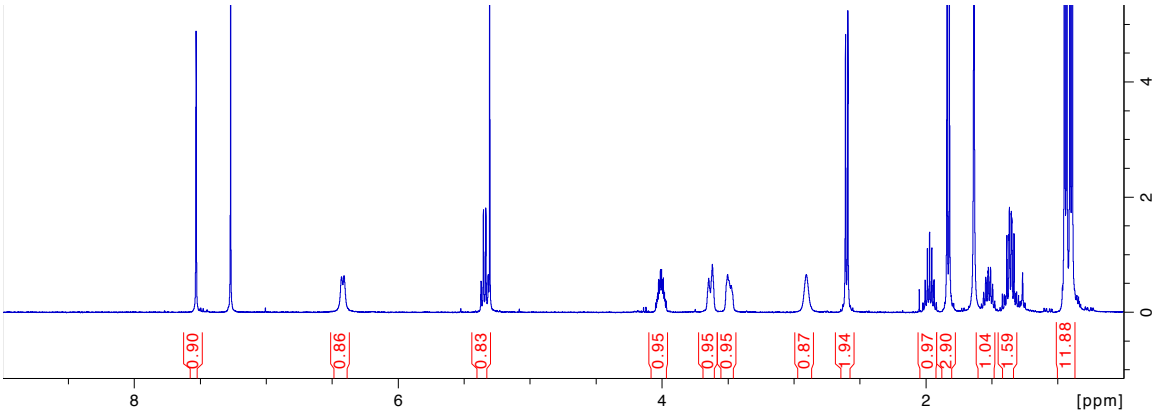
(*R*)-*N*-((*S*)-1-Hydroxy-4-methylpentan-2-yl)-2-(4-isobutyl-1*H*-1,2,3-triazol-1-yl)propanamide (DL-11lal)



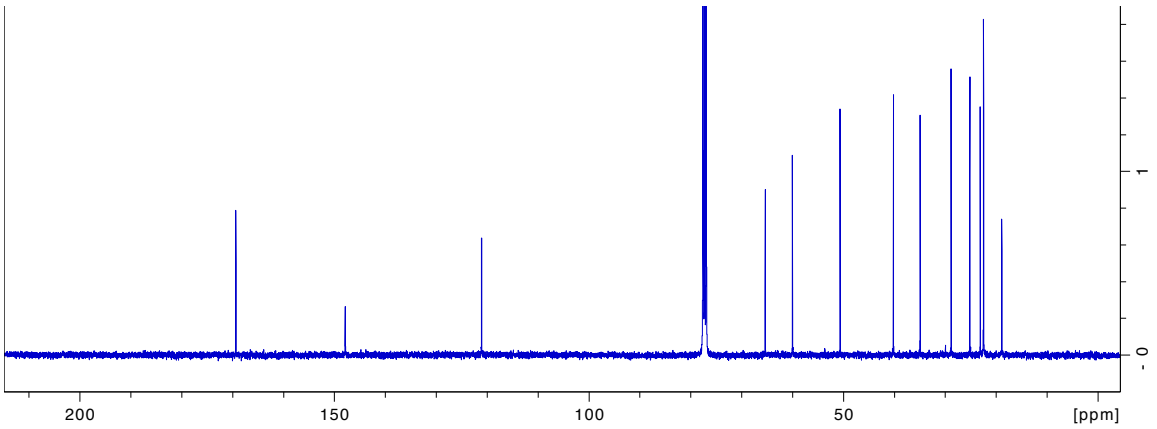
Data for DL-11lal solid, 62 % yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.53 (s, 1H), 6.42 (br, 1H), 5.34 (q, *J* = 7.2 Hz, 1H), 4.01 (m, 1H), 3.63 (m, 1H), 3.49 (m, 1H), 2.91 (br, 1H), 2.60 (d, *J* = 7.0 Hz, 2H), 1.97 (septet, *J* = 6.7 Hz, 1H), 1.83 (d, *J* = 7.2 Hz, 3H), 1.52 (m, 1H), 1.36 (m, 2H), 0.94 (d, *J* = 6.6 Hz, 6H), 0.90 (d, *J* = 6.6 Hz, 6H); ^{13}C NMR (100 MHz,

$CDCl_3$) δ 169.3, 147.8, 121.0, 65.3, 60.0, 50.6, 40.1, 34.9, 28.8, 25.1, 23.1, 22.5, 22.5, 22.4, 18.9; HRMS (ESI) m/z calcd for $C_{15}H_{29}N_4O_2^+$ 297.2285; found 297.2276 $(M+H)^+$

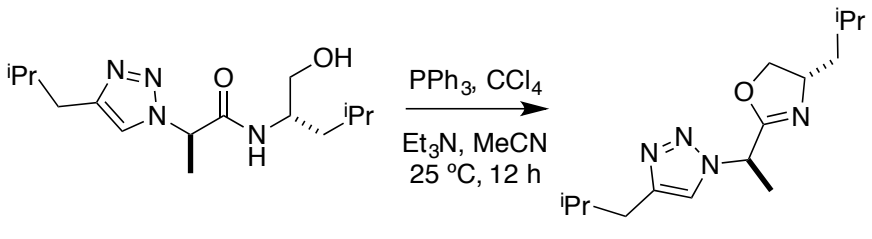
^{1}H NMR spectrum:



^{13}C NMR spectrum:



*Preparation of Compound DL-**3lal***

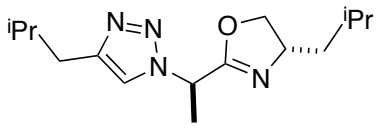


11 **3**

DL-**3lal**, 67%

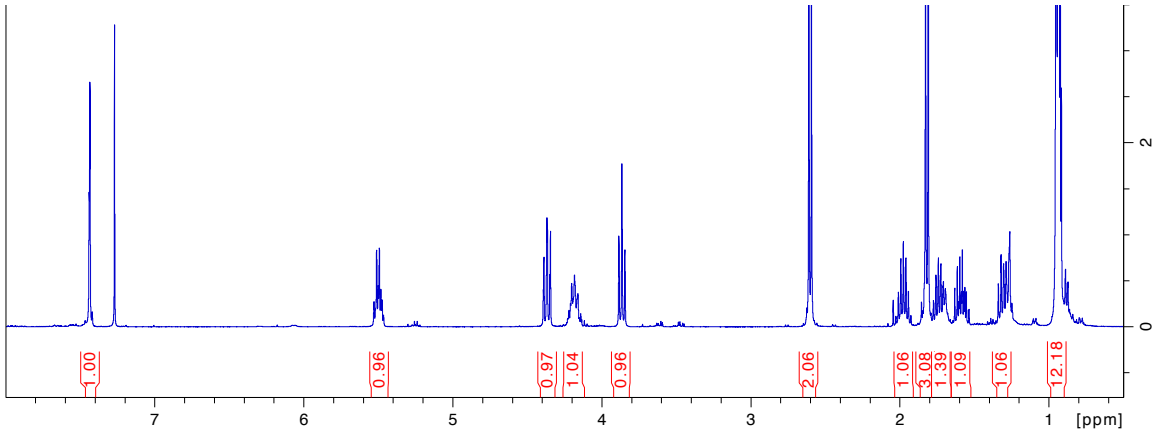
Compound DL-**11lal** (85 mg, 0.29 mmol) was dissolved in MeCN (3 mL), then PPh_3 (226 mg, 0.87 mmoL), Et_3N (0.16 mL, 1.15 mmol), and CCl_4 (0.15 mL, 1.58 mmol) were added subsequently. The reaction was stirred at room temperature for 12 h. Reaction was diluted with CH_2Cl_2, and then extracted with water twice and brine. Crude product was purified by flash column chromatography using EtOAc:hexanes (1:2) to obtain product as a solid (54 mg, 67%).

(*S*)-4-Isobutyl-2-((*R*)-1-(4-isobutyl-1*H*-1,2,3-triazol-1-yl)ethyl)-4,5-dihydrooxazole (DL-3lal)

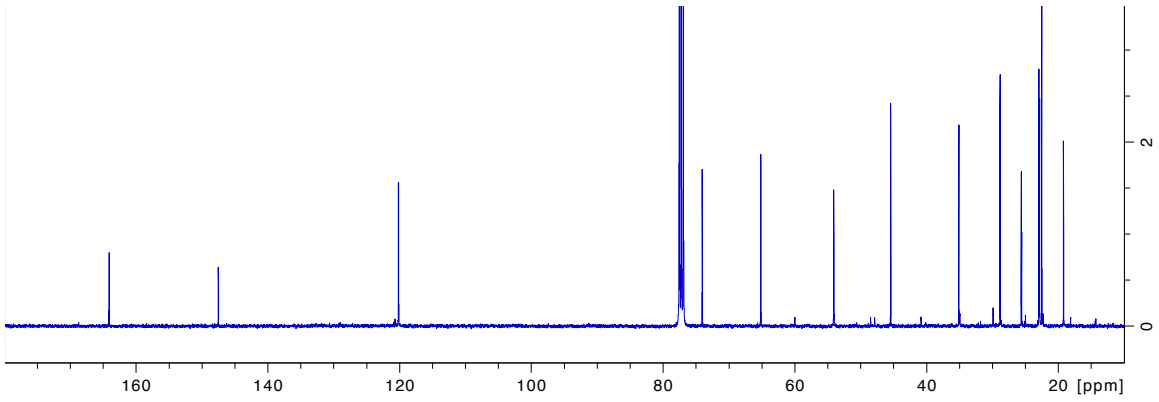


Data for DL-3lal solid, 62 % yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.44 (s, 1H), 5.50 (q, *J* = 7.1 Hz, 1H), 4.37 (m, 1H), 4.18 (m, 1H), 3.87 (t, *J* = 8.1 Hz, 1H), 2.60 (d, *J* = 7.0 Hz, 2H), 1.98 (septet, *J* = 6.7 Hz, 1H), 1.82 (d, *J* = 7.2 Hz, 3H), 1.73 (m, 1H), 1.59 (m, 1H), 1.30 (m, 1H), 0.94 (m, 12H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 164.1, 147.5, 120.2, 74.1, 65.2, 54.1, 45.4, 35.1, 28.8, 25.6, 23.0, 22.8, 22.5, 19.2; HRMS (ESI) m/z calcd for $C_{15}H_{27}N_4O^+$ 279.2179; found 279.2198 $(M+H)^+$

^{1}H NMR spectrum:



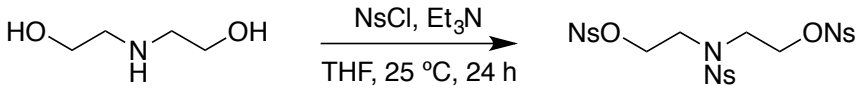
^{13}C NMR spectrum:



APPENDIX C

EXPERIMENTAL PROCEDURES FOR CHAPTER III

*Synthesis of Trisubstituted-nosyl Diethanolamine **13***

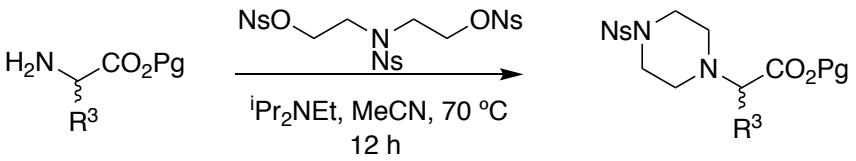


13

87%

Diethanolamine (5.6 g, 53 mmol) and Et_3N (27 mL, 192 mmol) were dissolved in 80 mL THF at 0 ºC under Ar atmosphere. The solution of *p*-nitrophenyl sulfonyl chloride (39.6 g, 179 mmol) in 45 mL THF was added dropwise over 30 minutes. Reaction was maintained at 0 ºC for 1 h, then allowed to stir at 25 ºC for 48 h. Solvent was completely removed by vacuo, and CH_2Cl_2 was added to dissolve the crude material. The mixture was kept in fridge to initiate precipitation, and the solid was collected by filtration. Remaining mother liquor was evaporated and reprecipitate again until no further solid forms. All crops were combined and recrystallized in hot THF/MeOH solution to obtain the final product as white solid (19.3 g, 55%). NMR spectra of this compound are identical to the reference.[168]

*General Procedure of Carboxylic-protected Piperazine-amino Acids **14** Syntheses*

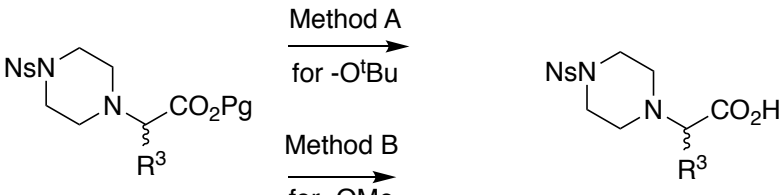


14

D-Asn-OMe, 26%; L-Asn-O^{t}Bu, 56%
L-Asn(Trt)-OMe, 93%; D-Asp(^{t}Bu)-OMe, 90%
L-Asp(tBu)-OMe, 65%; L-His(Trt)-OMe, 32%
L-Ile-OBn, 100%; L-Leu-O^{t}Bu, 48%
D-Lys(Boc)-OMe, 70%; L-Lys(Boc)-OMe, 81%
L-Phe-O^{t}Bu, 90%; D-Trp-OMe, 97%
L-Trp(Boc)-OMe, 66%; L-Val-O^{t}Bu, 46%

Trisubstituted-nosyl diethanolamine (1 eq) and amino-acid ester (1.2 eq) were dissolved in MeCN at 0.1 M concentration. Reaction was initiated by adding ^{i}Pr$_2$NEt (3 eq) and stirred at 70 ºC for 24 h under Ar atmosphere. After reaction completed, solvent was evaporated, and then redissolved in CH_2Cl_2. Mixture was extracted 3 times with 10% citric acid, 3 times with Sat. $NaHCO_3$ solution and brine. The organic layer was collected, dried with $MgSO_4$ and evaporated. Product was purified by flash column chromatography to obtain product as solid.

*General Procedure of Ester Deprotection **15***



14 **15**

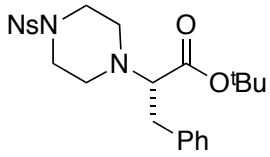
L-Asn, 99%; L-Asn(Trt), 97%
D-Asp(^{t}Bu), 95% ;L-Asp(tBu), 86%
L-His(Trt), 93%; L-Ile, 90%
D-Leu, 93%; L-Leu, 93%
D-Lys(Boc), 100% ; L-Lys(Boc), 100%
L-Phe, 89%; D-Trp, 88%

Method A (*tert*-butyl ester deprotection): Compounds **14** were stirred in 75% of TFA/CH_2Cl_2 for 3 h. After reaction completed, solvent was evaporated and repeated dissolvation/evaporation cycle with CH_2Cl_2 for 3 times. Dried crude was stirred in 1M HCl in MeOH for 15 min, then completely evaporated solvent to obtain products as HCl salts.

Method B (methyl ester deprotection): Compounds **14** (1 eq) and LiOH (3 eq) were dissolved in THF/H_2O mixture. MeOH was added dropwise until the solution became homogeneous. Mixture was stirred for 3 h until the reaction completed monitored by TLC. Organic solvent was removed by rotavap, and the aqueous layer was acidified with citric acid until pH = 2. The aqueous layer was extracted 4 times with ethyl acetate. Combined organic phase was extracted with brine, dried with $MgSO_4$ and evaporated to obtain product as solids.

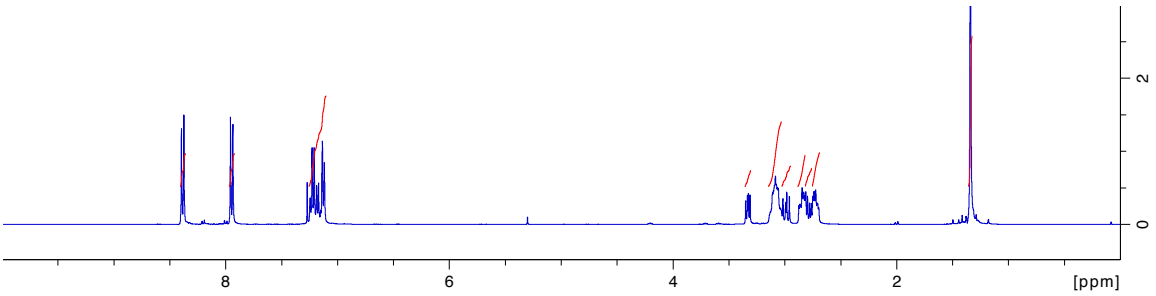
Method C (benzyl ester deprotection): Compound L-**14i** (1 eq) was dissolved in CH_2Cl_2 in the ice-propanol bath (-10 ºC) under Ar. Boron tribromide solution (3 eq) was added dropwise over 10 minutes, then the mixture was allowed to stir at 25 ºC for 4 h. The solution was then evaporated to dryness, redissolved crude material in 5 % Na_2CO_3 and extracted 3 times with diethyl ether. The aqueous layer was acidified with concentrated HCl to pH = 2, then extracted 3 times with ethyl acetate. Organic phase was washed with brine and dried with $MgSO_4$. Solvent was completely removed to obtain L-**15i** as solid.

***tert*-Butyl (*S*)-2-(4-((4-nitrophenyl)sulfonyl)piperazin-1-yl)-3-phenylpropanoate (L-14f)**

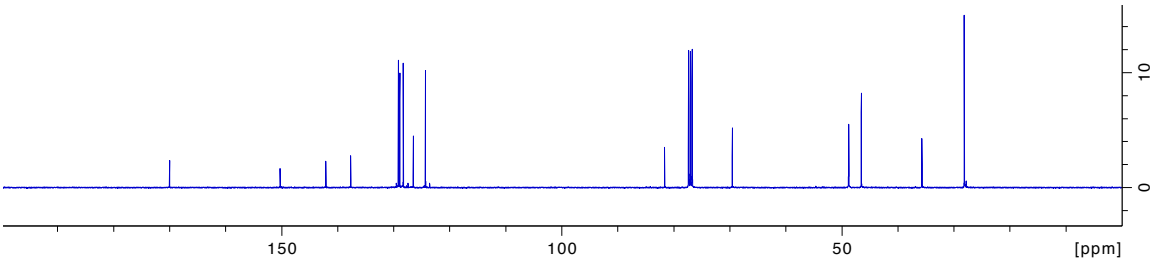


Data for L-14f. Yield 90%. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.38 (d, *J* = 8.8 Hz, 2H), 7.94 (d, *J* = 8.8 Hz, 2H), 7.26-7.09 (m, 5H), 3.33 (dd, *J* = 9.3, 6.0 Hz, 1H), 3.15-3.02 (m, 4H), 3.03-2.95 (m, 1H), 2.88-2.82 (m, 2H), 2.82-2.76 (m, 1H), 2.76-2.69 (m, 2H), 1.34 (s, 9H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 170.0, 150.2, 142.1, 137.6, 129.1, 128.9, 128.3, 126.5, 124.3, 81.6, 69.5, 48.8, 46.5, 35.7, 28.1; HRMS (ESI) m/z calcd for $C_{23}H_{30}N_3O_6S^+$ 476.1850; found 476.1843 $(M+H)^+$

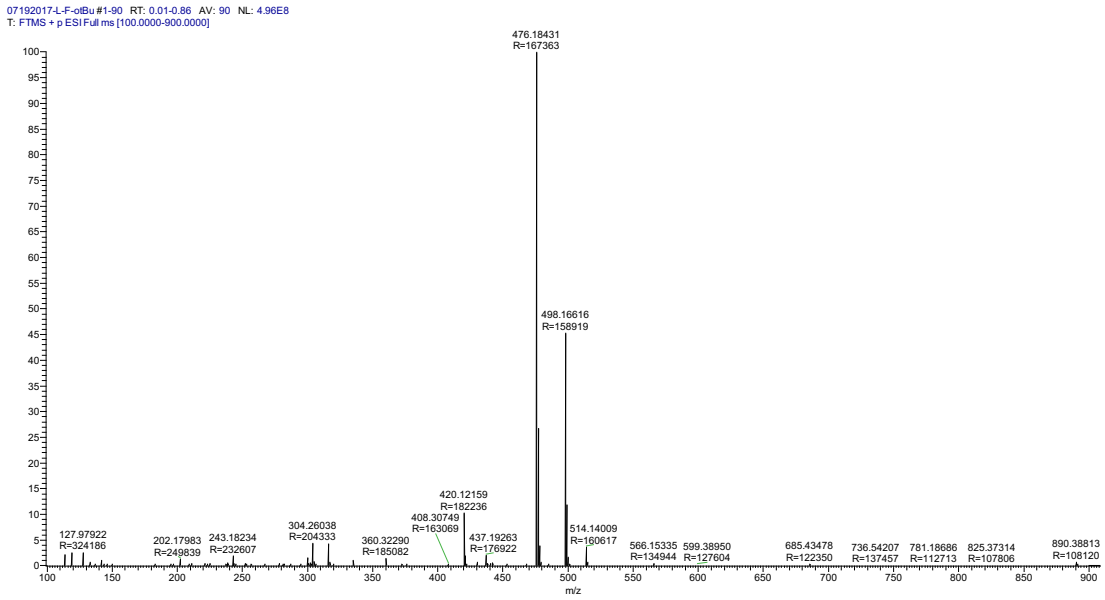
^{1}H NMR spectrum:



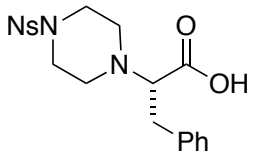
^{13}C NMR spectrum:



HRMS

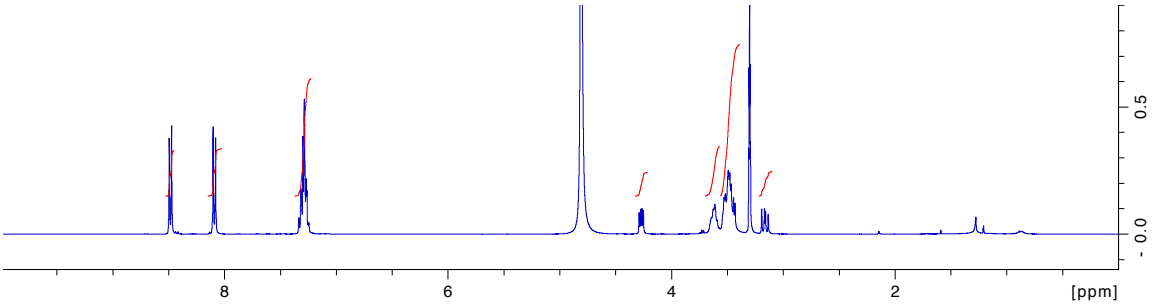


(*S*)-2-(4-((4-Nitrophenyl)sulfonyl)piperazin-1-yl)-3-phenylpropanoic acid (L-15f)

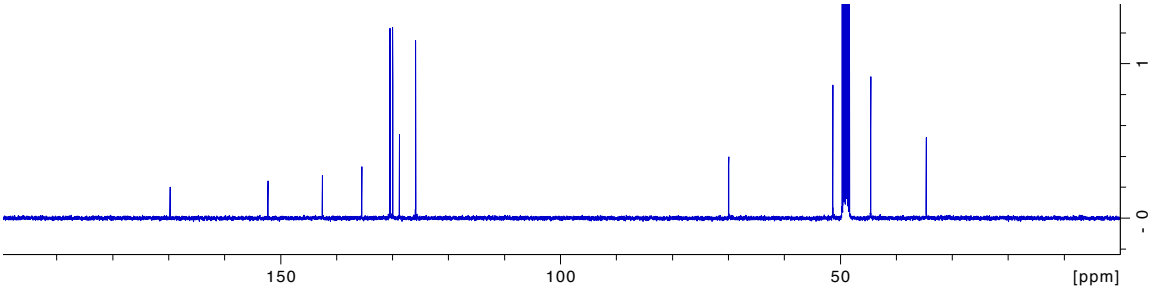


Data for L-15f. Yield 100%. ^{1}H NMR (400 MHz, MeOD-d_4) δ 8.48 (d, *J* = 8.9 Hz, 2H), 8.09 (d, *J* = 8.9 Hz, 2H), 7.35-7.24 (m, 5H), 4.27 (dd, *J* = 9.7, 5.1 Hz, 1H), 3.67-3.57 (m, 2H), 3.56-3.39 (m, 7H), 3.16 (dd, *J* = 13.6, 9.7 Hz, 1H); ^{13}C NMR (100 MHz, MeOD-d_4) δ 169.7, 152.2, 142.5, 135.5, 130.4, 130.2, 129.9, 128.7, 125.9, 69.9, 51.3, 44.5, 34.6; HRMS (ESI) m/z calcd for $C_{19}H_{22}N_3O_6S^+$ 420.1224; found 420.1219 $(M+H)^+$

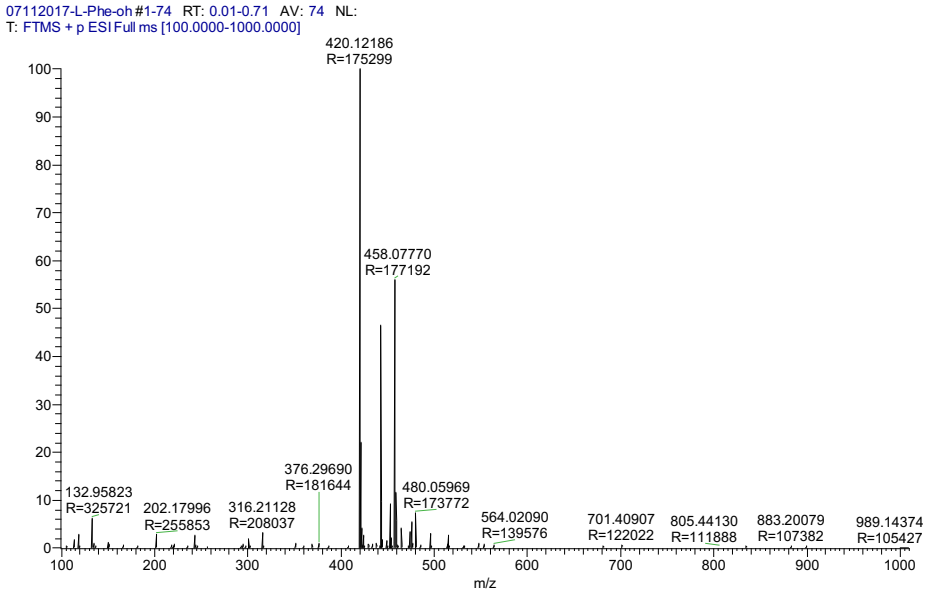
^{1}H NMR spectrum:



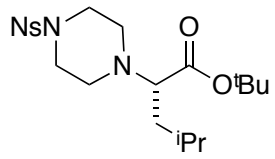
^{13}C NMR spectrum:



HRMS

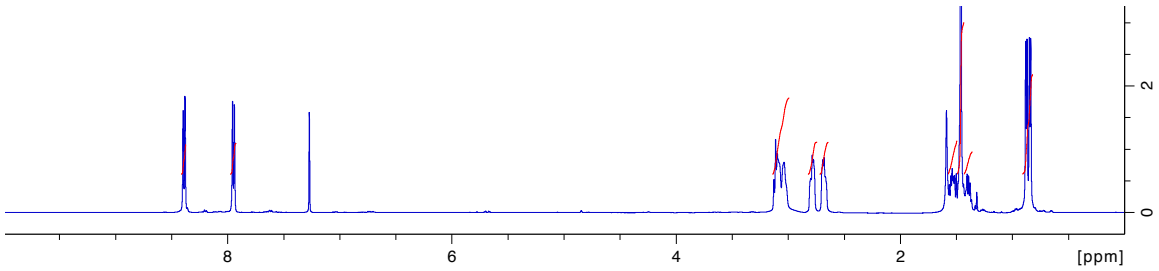


***tert*-Butyl (*S*)-4-methyl-2-(4-((4-nitrophenyl)sulfonyl)piperazin-1-yl)pentanoate (L-14l)**

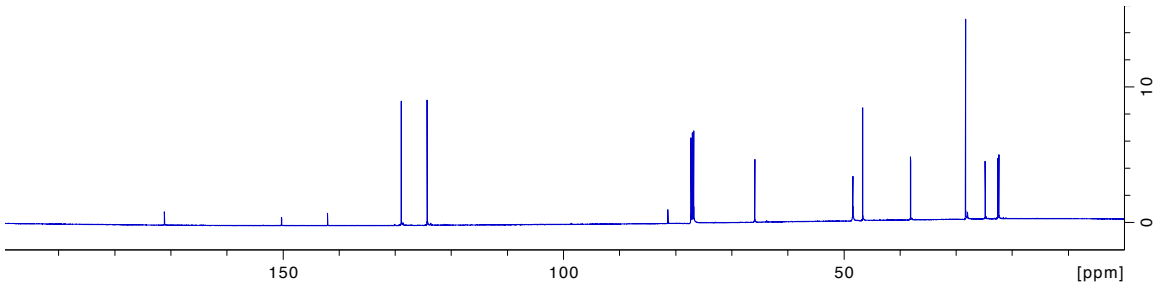


Data for L-14l. Yield 48%. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.39 (d, *J* = 7.9 Hz, 2H), 7.95 (d, *J* = 7.9 Hz, 2H), 3.14-2.98 (m, 5H), 2.79 (m, 2H), 2.68 (m, 2H), 1.57-1.49 (m, 2H), 1.46 (s, 9H), 1.43-1.35 (m, 1H), 0.86 (dd, *J* = 16.4, 5.8 Hz, 6H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 171.1, 150.2, 142.0, 128.9, 124.3, 81.4, 65.9, 48.4, 46.6, 38.1, 28.3, 24.8, 22.6, 22.4; HRMS (ESI) m/z calcd for $C_{20}H_{32}N_3O_6S^+$ 442.2006; found 442.2004 $(M+H)^+$

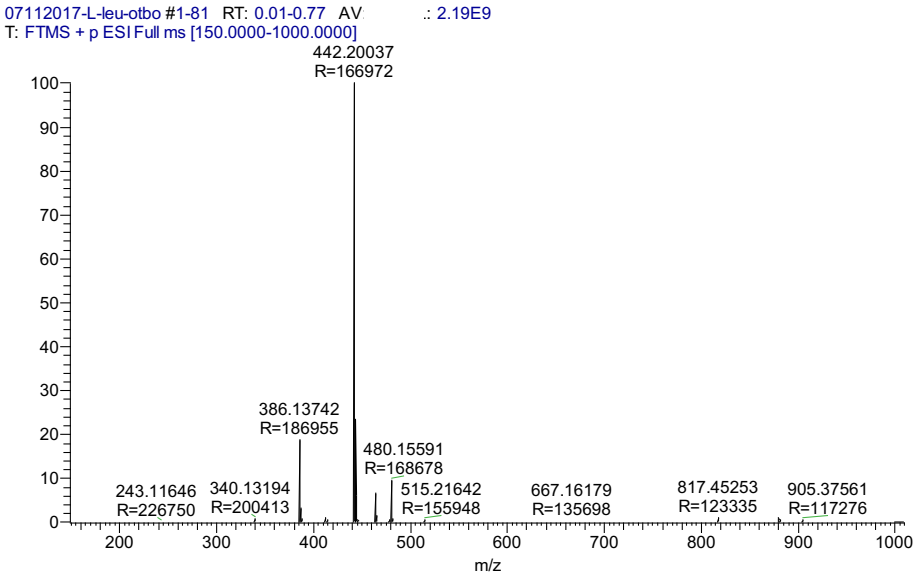
^{1}H NMR spectrum:



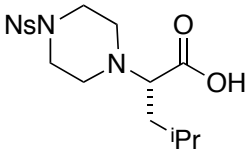
^{13}C NMR spectrum:



HRMS

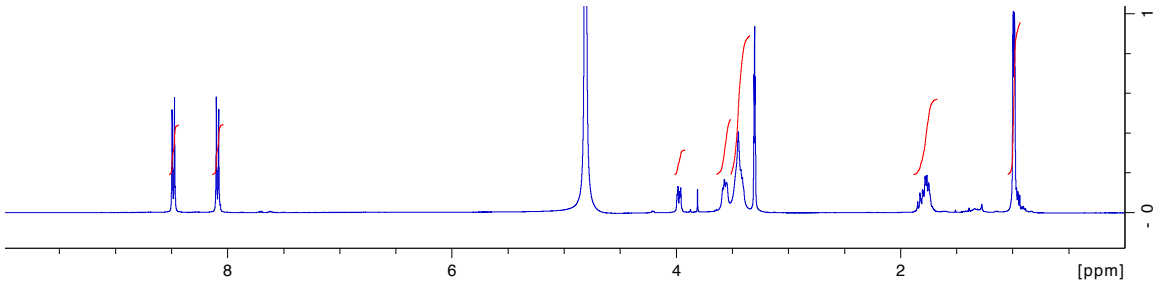


(*S*)-4-Methyl-2-(4-((4-nitrophenyl)sulfonyl)piperazin-1-yl)pentanoic acid (L-15l)

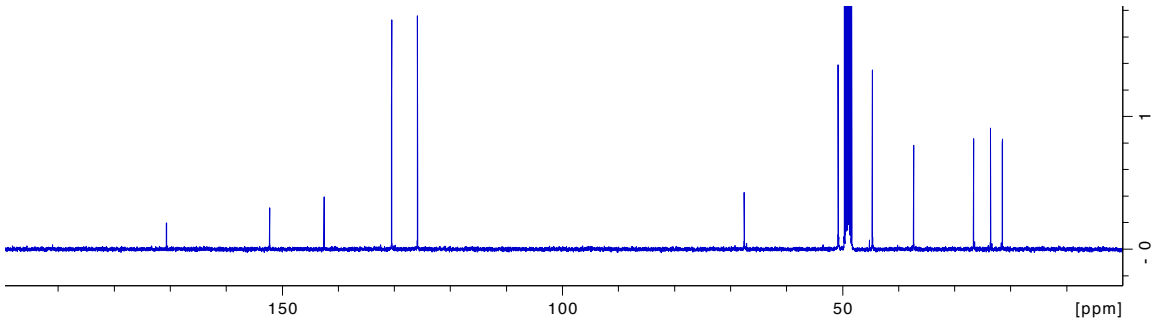


Data for L-15l. Yield 93%. ^{1}H NMR (400 MHz, MeOD-d_4) δ 8.48 (d, *J* = 8.9 Hz, 2H), 8.09 (d, *J* = 8.9 Hz, 2H), 3.97 (m, 1H), 3.57 (m, 2H), 3.44 (m, 6H), 1.88-1.66 (m, 3H), 0.99 (m, 6H); ^{13}C NMR (100 MHz, MeOD-d_4) δ 170.6, 152.2, 142.5, 130.5, 125.8, 67.5, 50.8, 44.7, 37.3, 26.6, 23.6, 21.5; HRMS (ESI) m/z calcd for $C_{16}H_{24}N_3O_6S^+$ 386.1380; found 386.1376 $(M+H)^+$

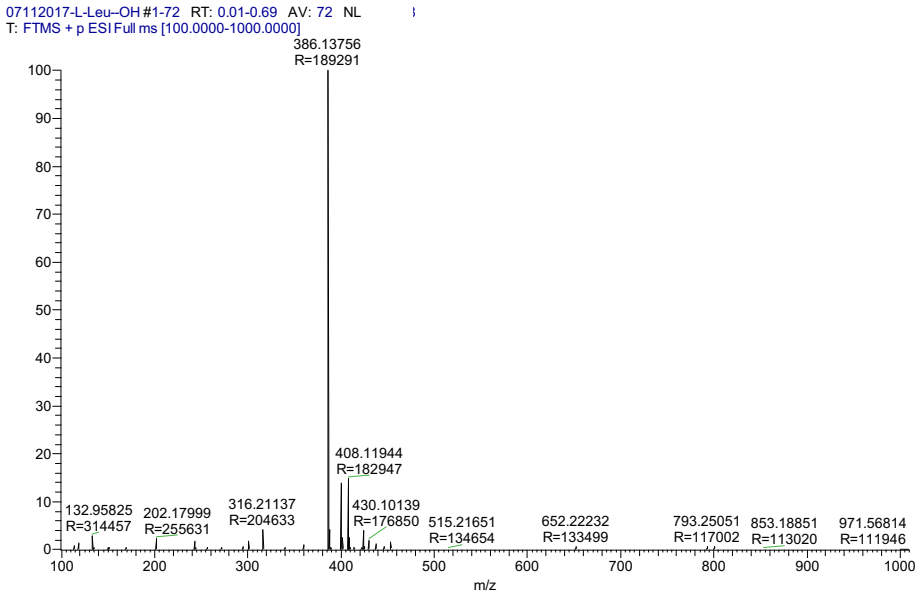
^{1}H NMR spectrum:



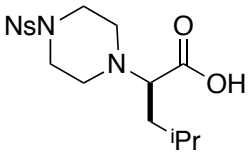
^{13}C NMR spectrum:



HRMS



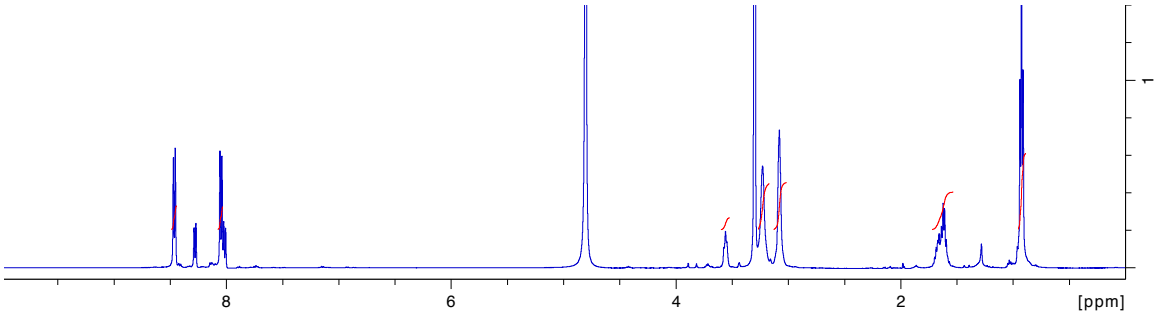
(*R*)-4-Methyl-2-(4-((4-nitrophenyl)sulfonyl)piperazin-1-yl)pentanoic acid (D-15l)



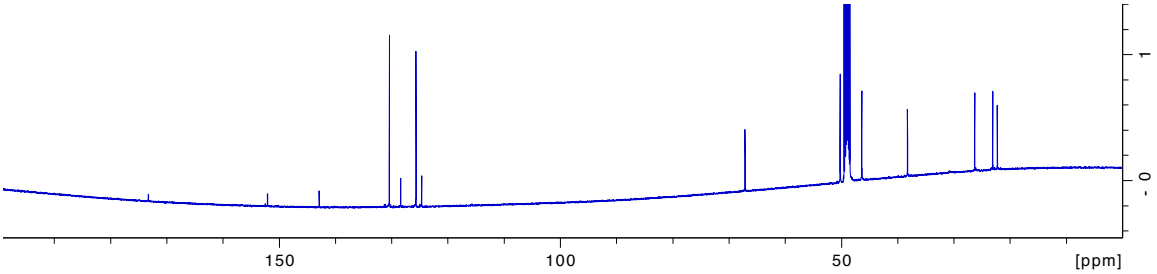
Data for D-15l. Yield 93%. ^{1}H NMR (400 MHz, MeOD-d_4) δ 8.46 (d, *J* = 8.0 Hz, 2H), 8.05 (d, *J* = 8.1 Hz, 2H), 3.56 (m, 1H), 3.23 (m, 4H), 3.08 (m, 4H), 1.72-1.56 (m, 3H), 0.92 (m, 6H); ^{13}C NMR (100 MHz, MeOD-d_4) δ 173.1, 152.3, 142.9, 130.4, 125.6, 67.1,

50.2, 46.4, 38.2, 26.3, 23.1, 22.3 HRMS (ESI) m/z calcd for $C_{16}H_{24}N_3O_6S^+$ 385.1380; found 386.1375 $(M+H)^+$

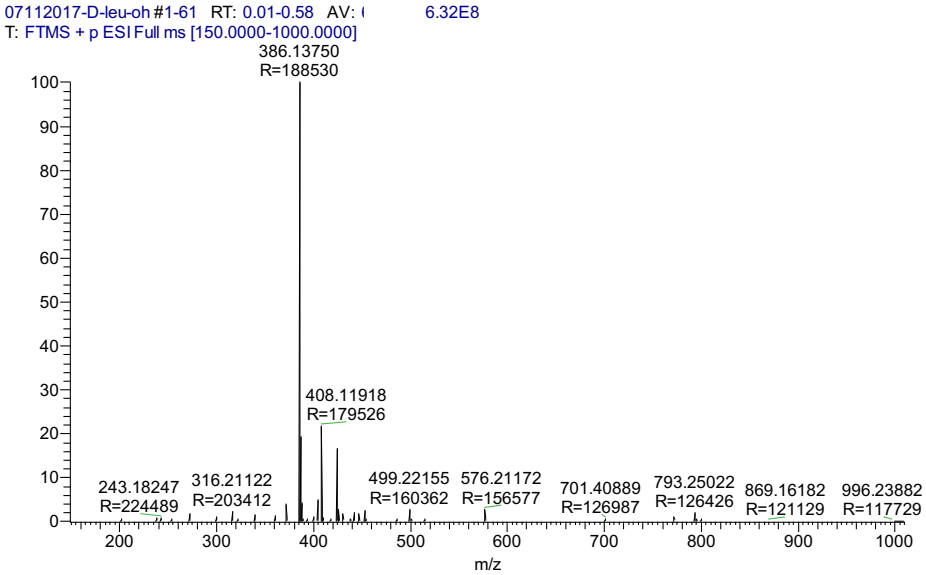
^{1}H NMR spectrum:



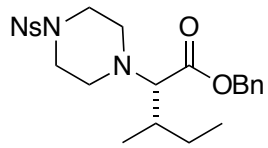
^{13}C NMR spectrum:



HRMS

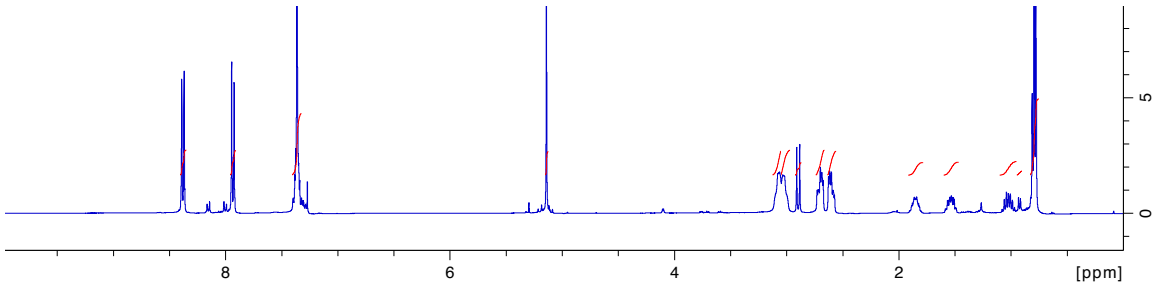


Benzyl (2*S*,3*S*)-3-methyl-2-(4-((4-nitrophenyl)sulfonyl)piperazin-1-yl)pentanoate (L-14i)

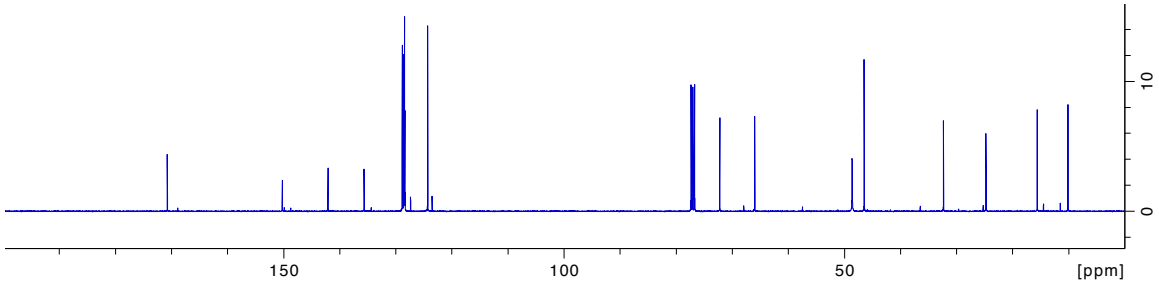


Data for L-14i. Yield 89%. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.38 (d, *J* = 8.9 Hz, 2H), 7.93 (d, *J* = 8.9 Hz, 2H), 7.36 (m, 5H), 5.14 (s, 2H), 3,13-2.97 (m, 4H), 2.90 (d, *J* = 10.6 Hz, 1H), 2.70 (m, 2H), 2.60 (m, 2H), 1.86 (m, 1H), 1.54 (m, 1H), 1.02 (m, 1H), 0.79 (m, 6H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 170.7, 150.2, 142.1, 135.6, 128.8, 128.6, 128.4, 128.3, 124.3, 72.2, 66.0, 48.6, 46.5, 32.3, 24.7, 15.6, 10.1; HRMS (ESI) m/z calcd for $C_{23}H_{30}N_3O_6S^+$ 476.1850; found 476.1846 $(M+H)^+$

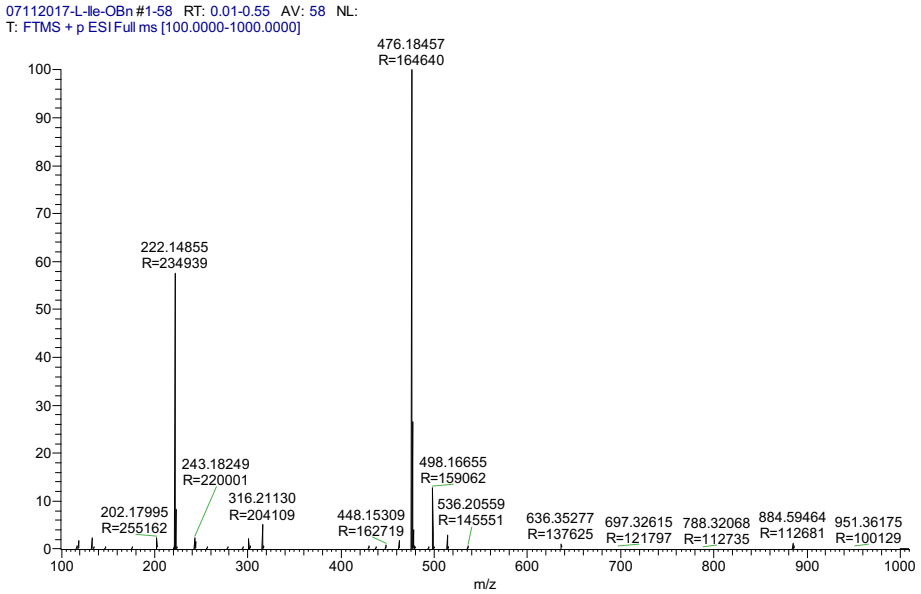
^{1}H NMR spectrum:



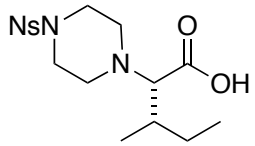
^{13}C NMR spectrum:



HRMS

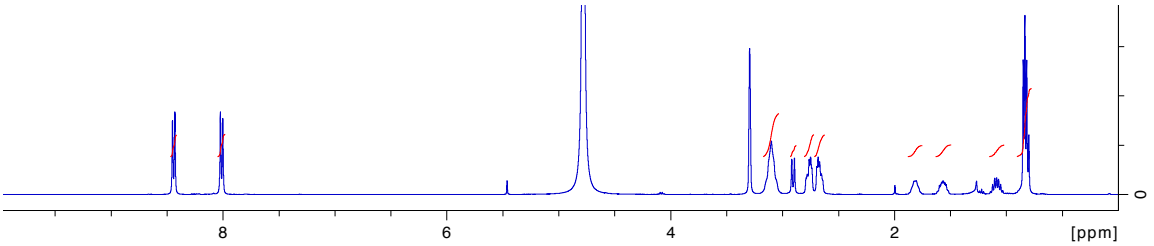


(2*S*,3*S*)-3-Methyl-2-(4-((4-nitrophenyl)sulfonyl)piperazin-1-yl)pentanoic acid (L-1i5)

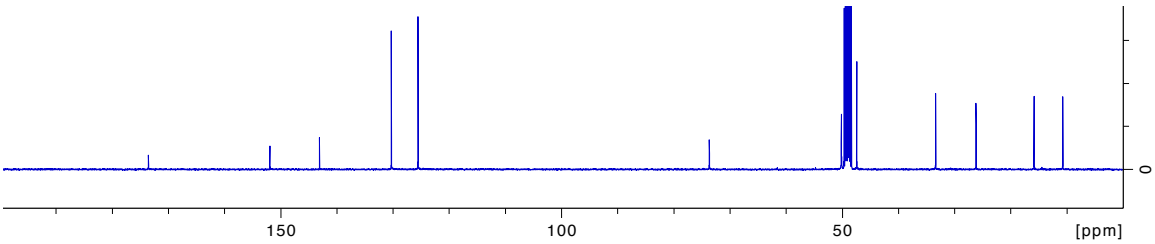


Data for L-15i. Yield 90%. ^{1}H NMR (400 MHz, MeOD-d_4) δ 8.44 (d, *J* = 8.6 Hz, 2H), 8.01 (d, *J* = 8.6 Hz, 2H), 3.10 (m, 4H), 2.90 (d, *J* = 9.6 Hz, 1H), 2.76 (m, 2H), 2.67 (m, 2H), 1.81 (m, 1H), 1.57 (m, 1H), 1.09 (m, 1H), 0.83 (m, 6H); ^{13}C NMR (100 MHz, MeOD-d_4) δ 173.5, 151.9, 143.1, 130.3, 125.5, 73.7, 50.1, 47.4, 33.4, 26.2, 15.8, 10.7; HRMS (ESI) m/z calcd for $C_{16}H_{23}N_3O_6S$ 386.1380; found 386.1376 $(M+H)^+$

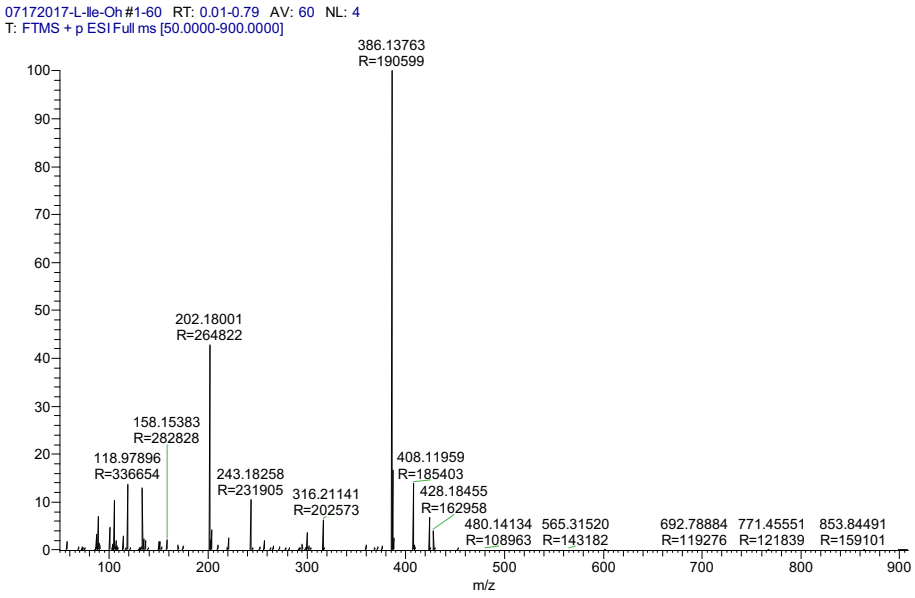
^{1}H NMR spectrum:



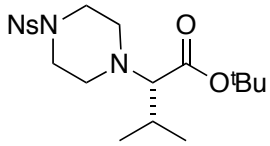
^{13}C NMR spectrum:



HRMS

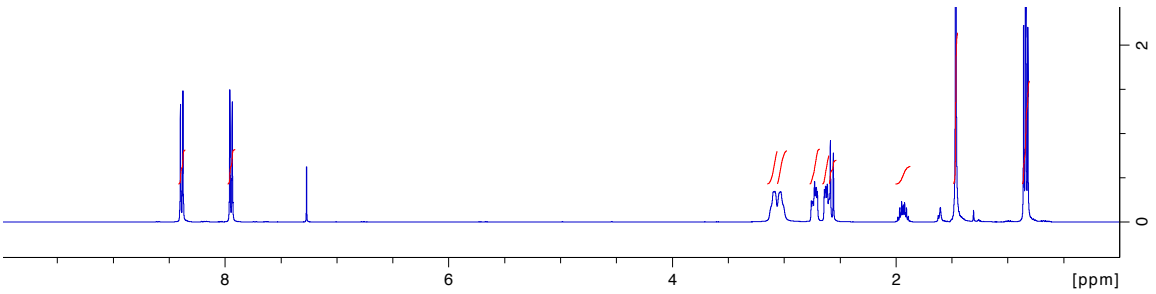


***tert*-Butyl (*S*)-3-methyl-2-(4-((4-nitrophenyl)sulfonyl)piperazin-1-yl)butanoate (L-14v)**

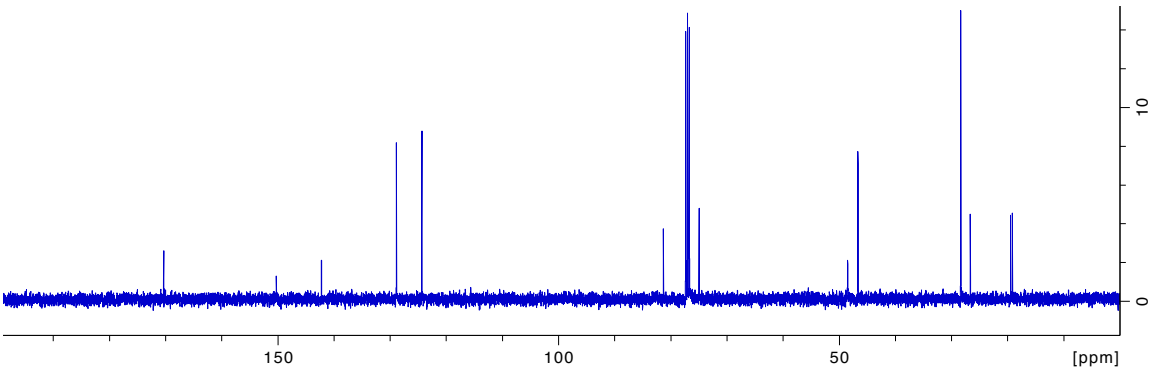


Data for L-14v. Yield 46%. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.39 (d, *J* = 8.9 Hz, 2H), 7.95 (d, *J* = 8.9 Hz, 2H), 3.14-2.97 (m, 4H), 2.73 (m, 2H), 2.61 (m, 2H), 2.57 (d, *J* = 10.7 Hz, 1H), 1.94 (m, 1H), 1.46 (s, 9H), 0.84 (dd, *J* = 8.4, 6.6 Hz, 6H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 170.3, 150.2, 142.2, 128.9, 124.3, 81.3, 77.3, 77.0, 76.7, 74.9, 48.5, 46.6, 28.4, 26.6, 19.5, 19.2; HRMS (ESI) m/z calcd for $C_{19}H_{30}N_3O_6S^+$ 428.1850; found 428.1843 $(M+H)^+$

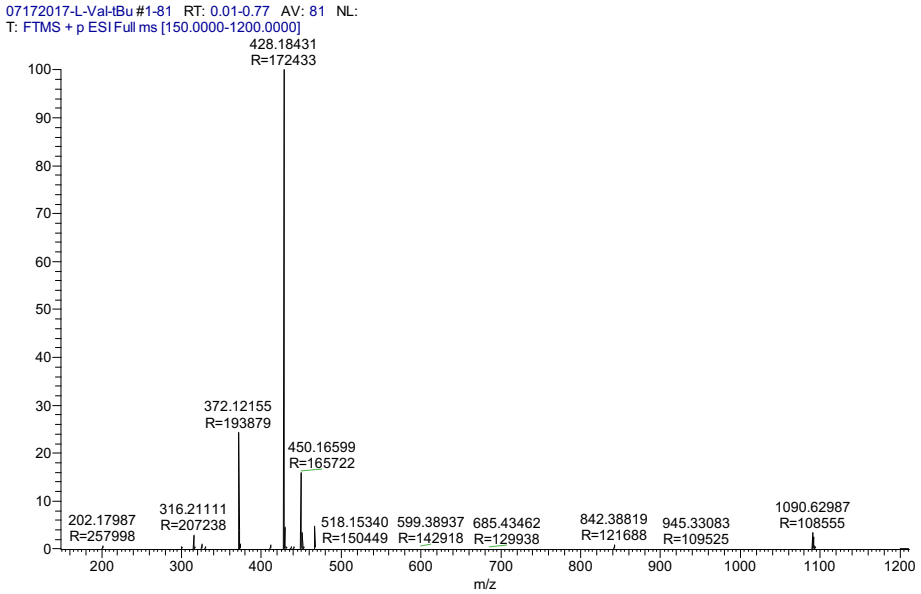
^{1}H NMR spectrum:



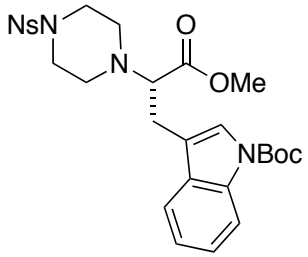
^{13}C NMR spectrum:



HRMS

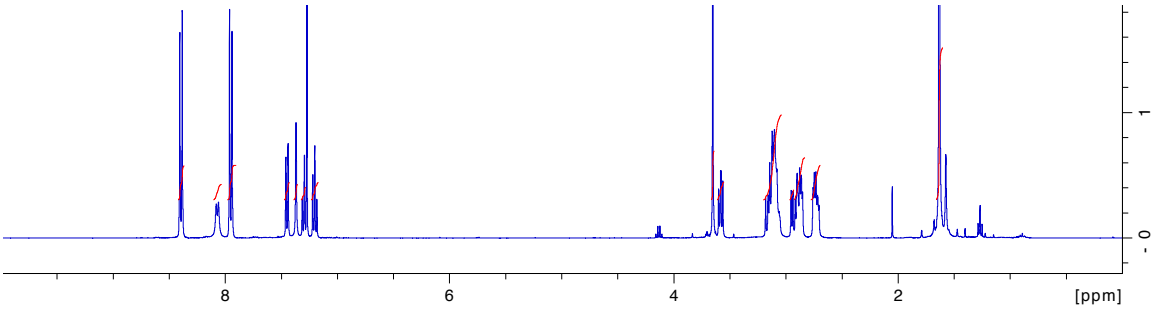


***tert*-Butyl (*S*)-3-(3-methoxy-2-(4-((4-nitrophenyl)sulfonyl)piperazin-1-yl)-3-oxopropyl)-1*H*-indole-1-carboxylate (L-14w')**

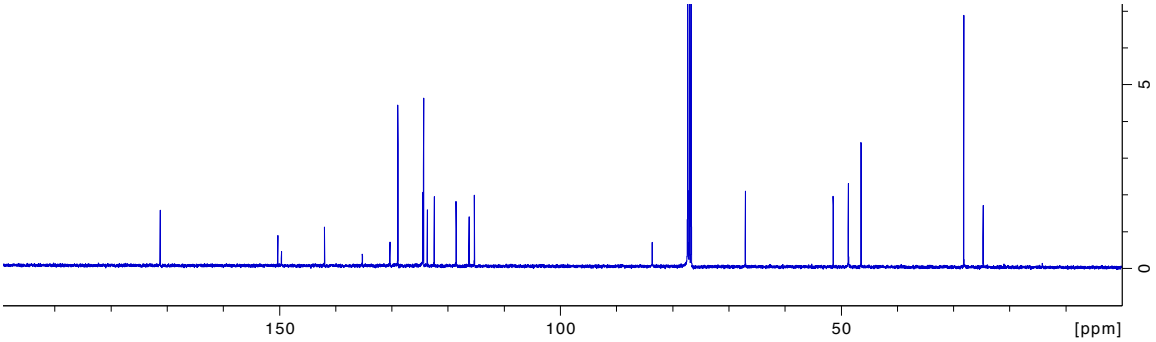


Data for L-14w'. Yield 66%. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.39 (d, *J* = 8.8 Hz, 2H), 8.06 (d, *J* = 7.9 Hz, 1H), 7.95 (d, *J* = 8.8 Hz, 2H), 7.45 (d, *J* = 7.7 Hz, 1H), 7.45 (s, 1H), 7.30 (m, 1H), 7.20 (m, 1H), 3.65 (s, 3H), 3.58 (dd, *J* = 8.4, 6.5 Hz, 1H), 3.20-3.03 (m, 5H), 2.96-2.83 (m, 3H), 2.78-2.69 (m, 2H), 1.63 (s, 9H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 171.2, 150.3, 149.6, 142.0, 135.2, 130.3, 128.9, 124.4, 124.3, 123.6, 122.4, 118.5, 116.2, 115.3, 83.6, 67.0, 51.4, 48.7, 46.4, 28.2, 24.7; HRMS (ESI) m/z calcd for $C_{27}H_{33}N_4O_8S^+$ 573.2014; found 573.2011 $(M+H)^+$

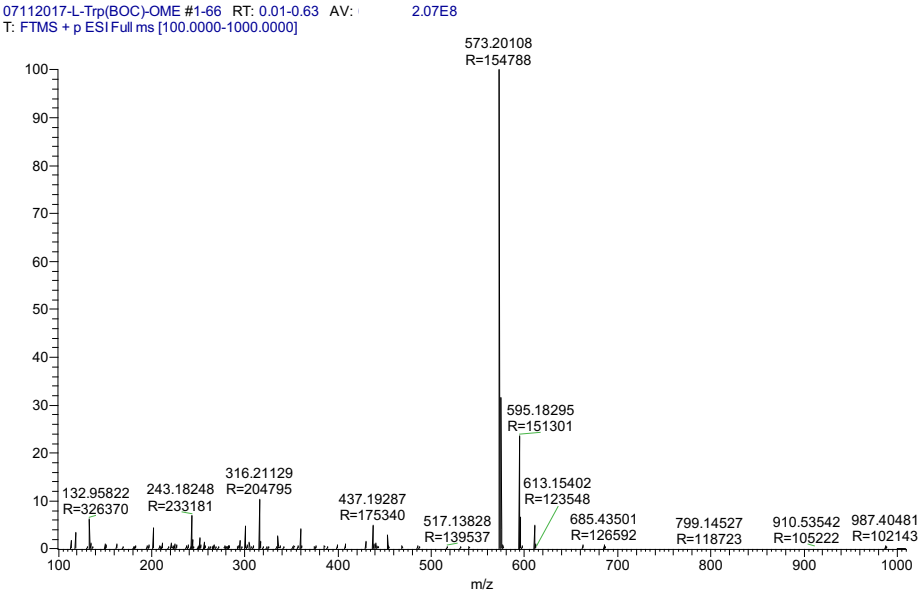
^{1}H NMR spectrum:



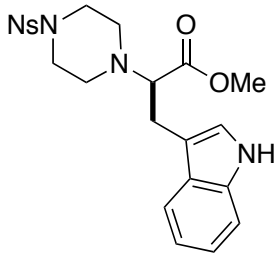
^{13}C NMR spectrum



HRMS

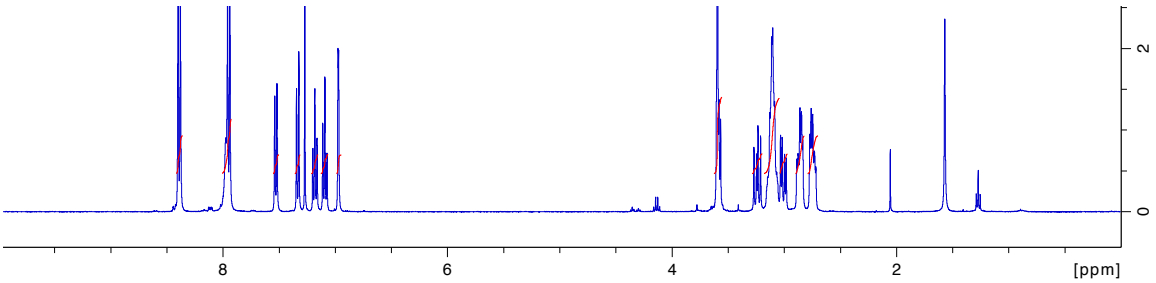


Methyl (*R*)-3-(1*H*-indol-3-yl)-2-(4-((4-nitrophenyl)sulfonyl)piperazin-1-yl)propanoate (D-14w)

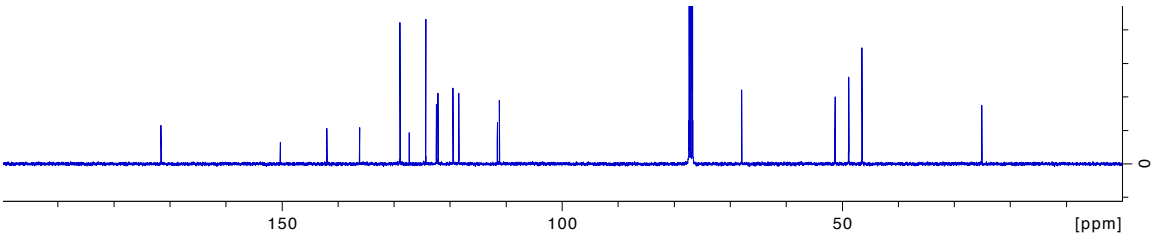


Data for D-14w. Yield 97%. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.39 (d, *J* = 8.8 Hz, 2H), 7.98 (br, 1H), 7.95 (d, *J* = 8.8 Hz, 2H), 7.53 (d, *J* = 7.8 Hz, 1H), 7.33 (d, *J* = 8.1 Hz, 1H), 7.18 (t, *J* = 7.5 Hz, 1H), 7.09 (t, *J* = 7.5 Hz, 1H), 6.97 (s, 1H), 3.60 (s, 3H), 3.57 (d, *J* = 5.7 Hz, 1H), 3.24 (dd, *J* = 14.4, 9.3 Hz, 1H), 3.17-3.05 (m, 4H), 3.01 (dd, *J* = 14.4, 5.6 Hz, 1H), 2.87 (m, 2H), 2.75 (m, 2H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 171.6, 150.3, 142.0, 136.1, 128.9, 127.3, 124.3, 122.4, 122.1, 119.5, 118.4, 111.5, 111.2, 67.9, 51.3, 48.8, 46.5, 25.1; HRMS (ESI) m/z calcd for $C_{22}H_{25}N_4O_6S^+$ 473.1489; found 473.1486 $(M+H)^+$

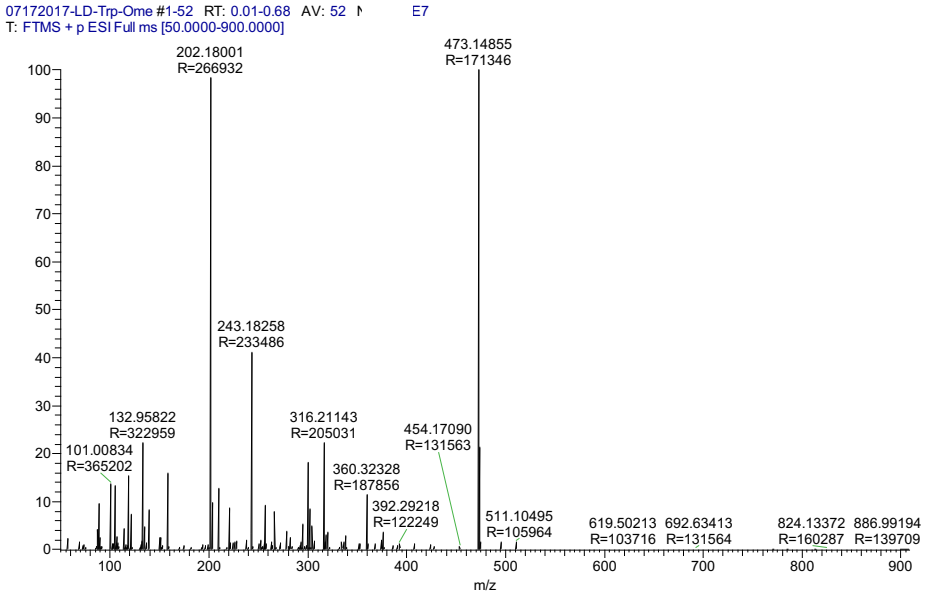
^{1}H NMR spectrum:



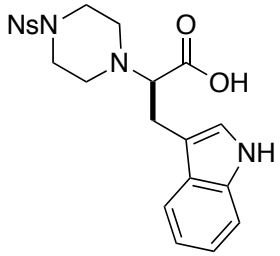
^{13}C NMR spectrum:



HRMS

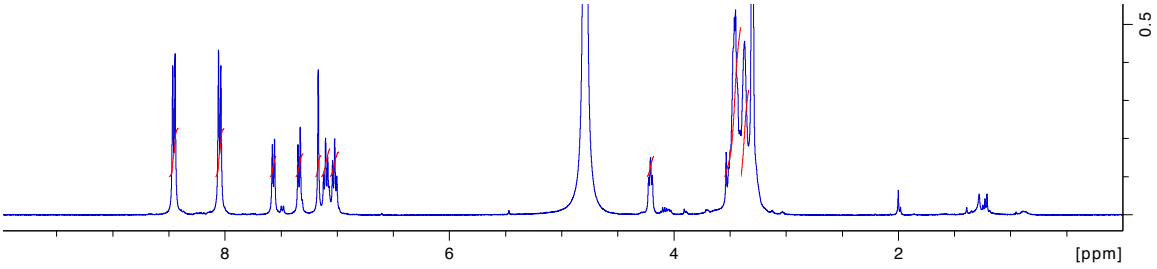


(*R*)-3-(1*H*-Indol-3-yl)-2-(4-((4-nitrophenyl)sulfonyl)piperazin-1-yl)propanoic acid (D-15w)

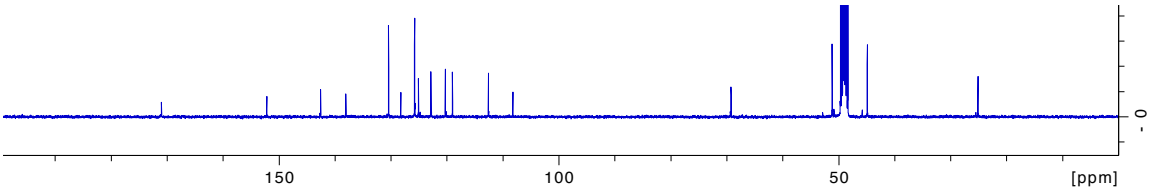


Data for D-15w. Yield 88%. ^{1}H NMR (400 MHz, MeOD-d_4) δ 8.45 (d, *J* = 8.6 Hz, 2H), 8.05 (d, *J* = 8.6 Hz, 2H), 7.57 (d, *J* = 7.8 Hz, 1H), 7.34 (d, *J* = 8.0 Hz, 1H), 7.17 (s, 1H), 7.10 (t, *J* = 7.5 Hz, 1H), 7.02 (t, *J* = 7.5 Hz, 1H), 4.21 (m, 1H), 3.55-3.40 (m, 6H), 3.40-3.33 (m, 4H); ^{13}C NMR (100 MHz, MeOD-d_4) δ 171.0, 152.2, 142.6, 138.1, 130.4, 128.2, 125.8, 125.1, 122.9, 120.3, 119.0, 112.6, 108.2, 69.2, 51.2, 44.9, 25.1; HRMS (ESI) m/z calcd for $C_{21}H_{23}N_4O_6S^+$ 459.1333; found 459.1329 $(M+H)^+$

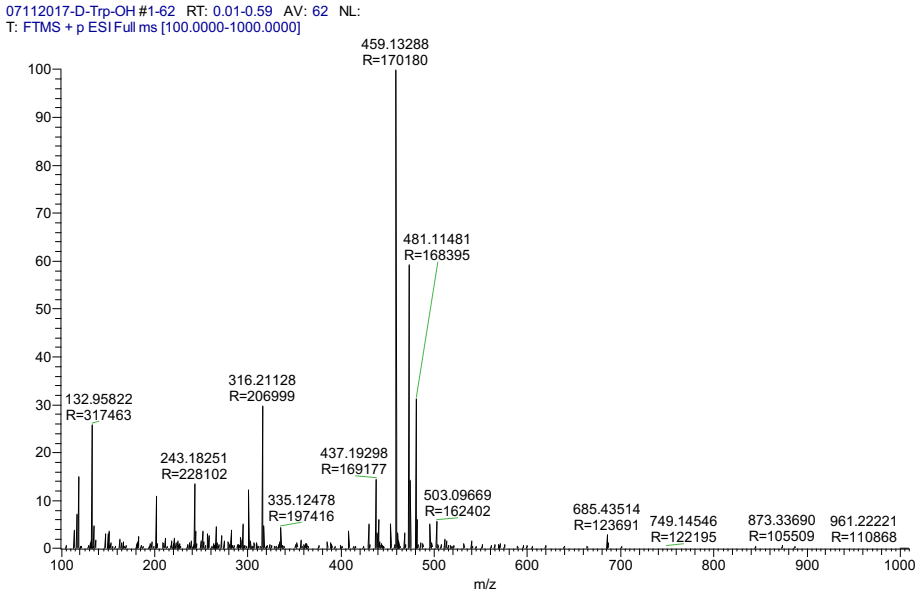
^{1}H NMR spectrum:



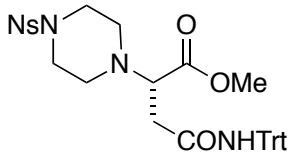
^{13}C NMR spectrum:



HRMS

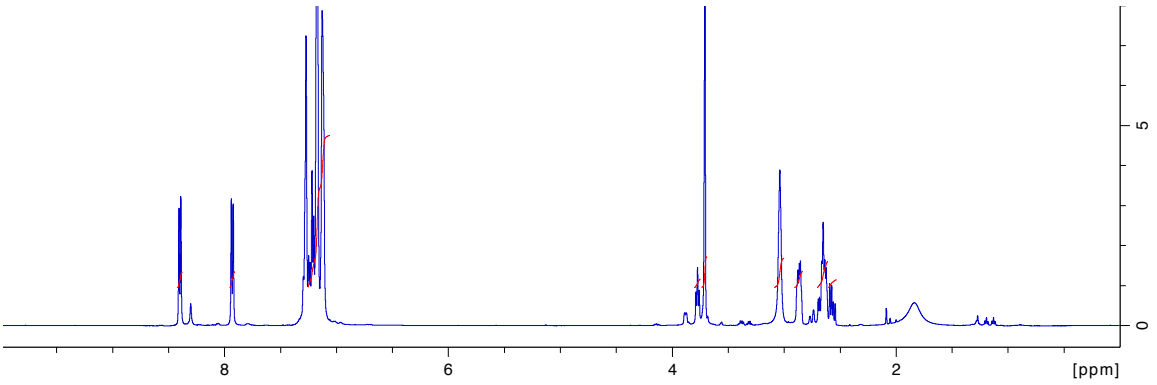


Methyl (*S*)-2-(4-((4-nitrophenyl)sulfonyl)piperazin-1-yl)-4-oxo-4-(tritylamino)butanoate (L-14n')

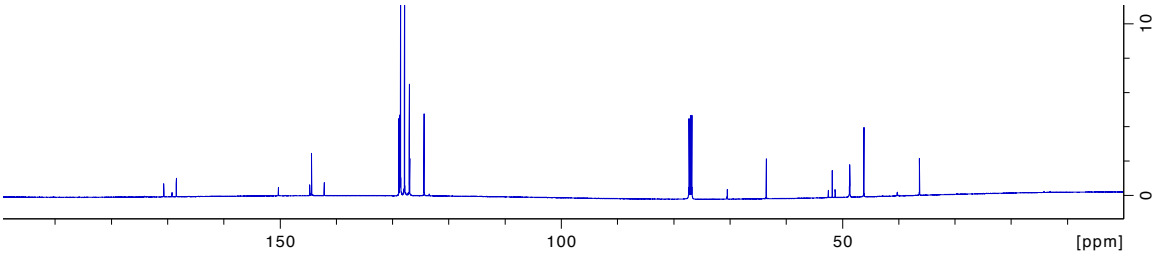


Data for L-14n'. Yield 93%. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.40 (d, *J* = 7.8 Hz, 2H), 7.93 (d, *J* = 7.8 Hz, 2H), 7.25-7.09 (m, 15H), 3.77 (t, *J* = 7.1 Hz, 1H), 3.71 (s, 4H), 3.04 (m, 4H), 2.87 (m, 2H), 2.70-2.61 (m, 3H), 2.57 (dd, *J* = 15.5, 8.4 Hz, 1H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 170.7, 169.2, 168.4, 150.3, 144.7, 144.4, 142.1, 128.8, 128.7, 128.5, 127.8, 127.0, 126.9, 124.4, 70.4, 63.5, 51.8, 48.7, 46.2, 36.3; HRMS (ESI) m/z calcd for $C_{34}H_{35}N_4O_7S^+$ 643.2221; found 643.2219 $(M+H)^+$

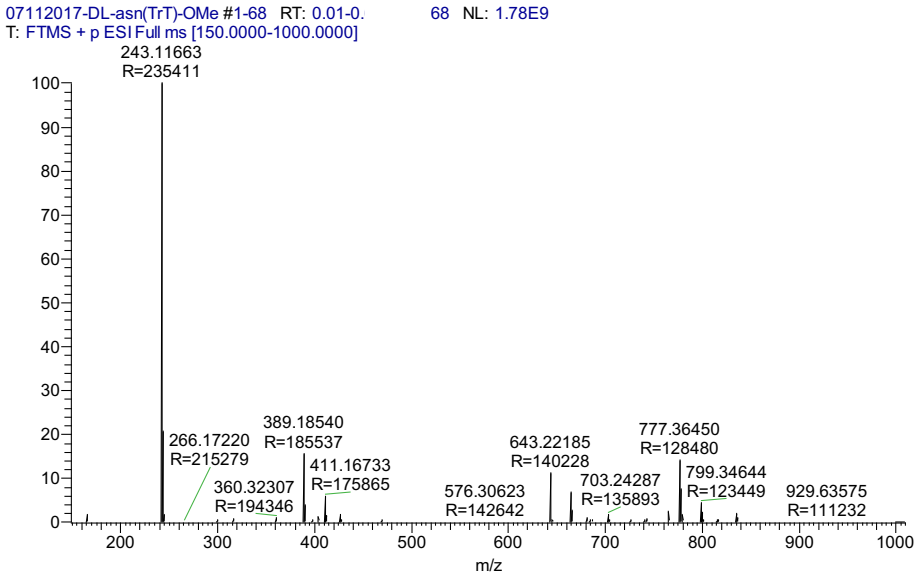
^{1}H NMR spectrum:



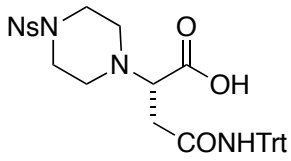
^{13}C NMR spectrum:



HRMS

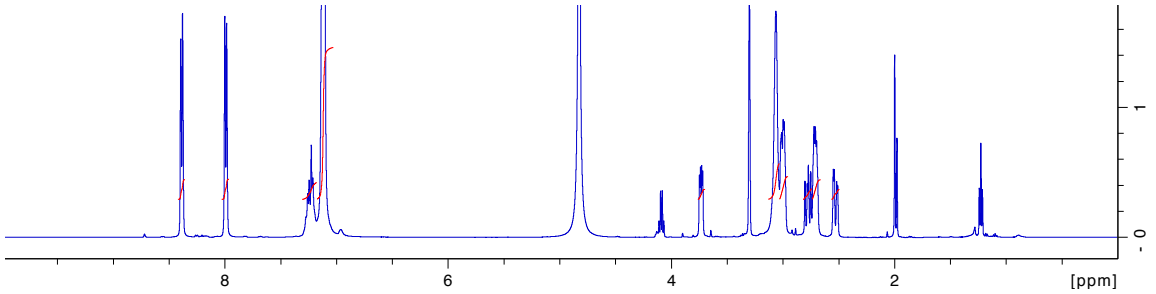


(*S*)-2-(4-((4-Nitrophenyl)sulfonyl)piperazin-1-yl)-4-oxo-4-(tritylamino)butanoic acid (L-15n')

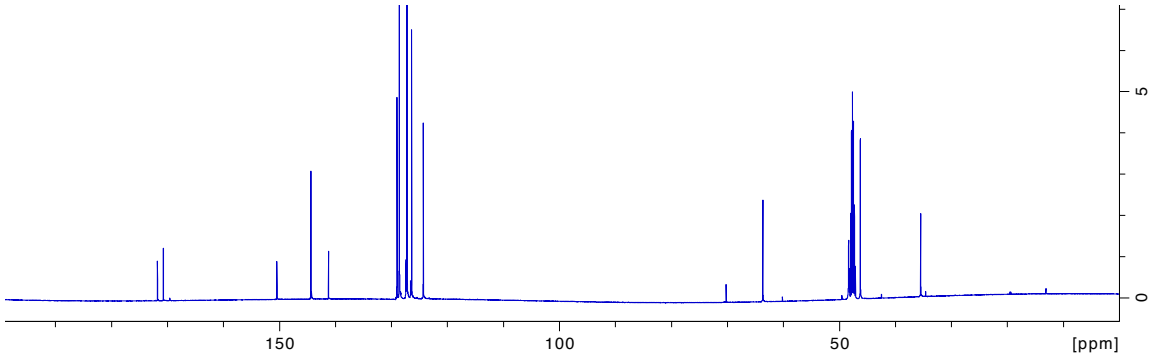


Data for L-15n'. Yield 97%. ^{1}H NMR (400 MHz, MeOD-d_4) δ 8.39 (d, *J* = 8.0 Hz, 2H), 7.99 (d, *J* = 8.1 Hz, 2H), 7.12 (m, 15H), 3.73 (dd, *J* = 10.0, 4.5 Hz, 1H), 3.06 (m, 4H), 3.00 (m, 2H), 2.78 (dd, *J* = 15.3, 10.4 Hz, 1H), 2.71 (m, 2H), 2.53 (dd, *J* = 15.4, 4.1 Hz, 2H); ^{13}C NMR (100 MHz, MeOD-d_4) δ 171.8, 170.7, 150.4, 144.4, 141.2, 129.0, 128.6, 127.2, 126.4, 124.3, 70.2, 63.6, 48.3, 46.2, 35.5; HRMS (ESI) m/z calcd for $C_{33}H_{33}N_4O_7S^+$ 629.2064; found 629.2056 $(M+H)^+$

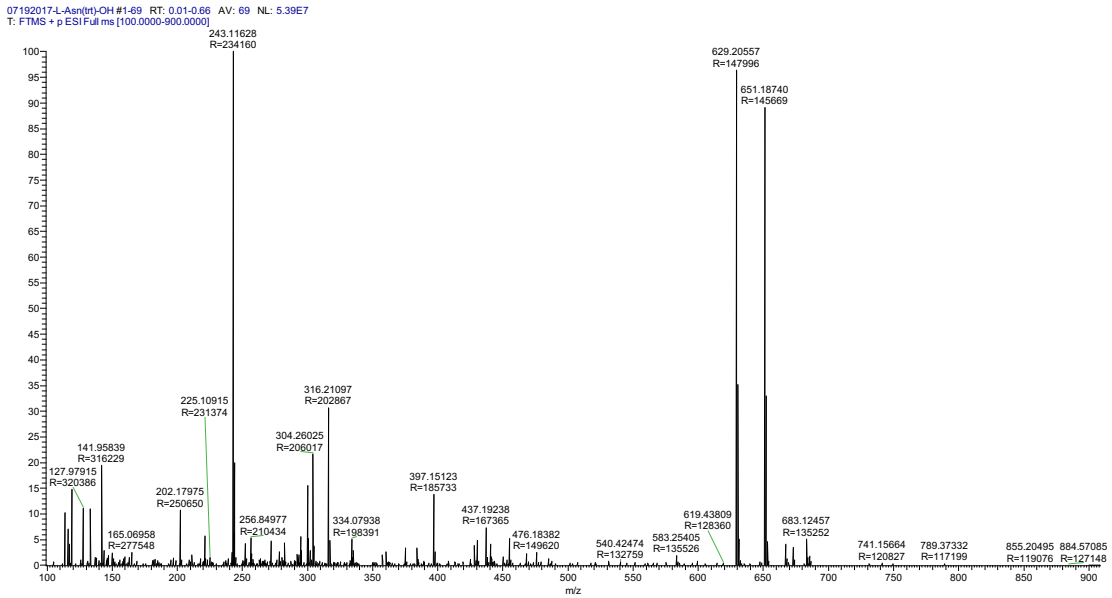
^{1}H NMR spectrum:



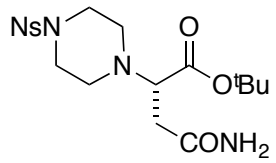
^{13}C NMR spectrum



HRMS

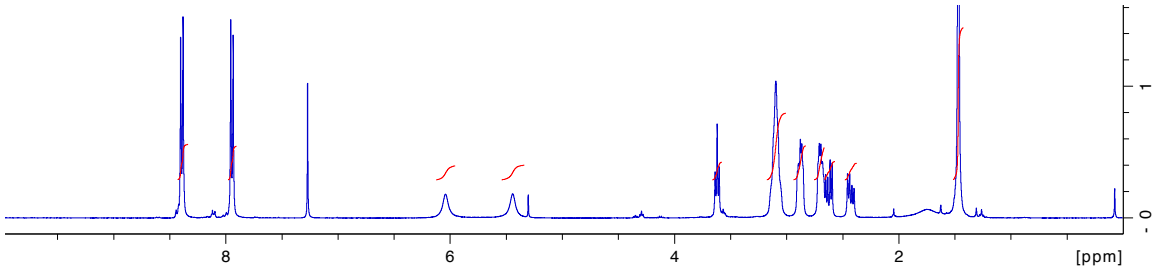


***tert*-Butyl (*S*)-4-amino-2-(4-((4-nitrophenyl)sulfonyl)piperazin-1-yl)-4-oxobutanoate (L-14n)**

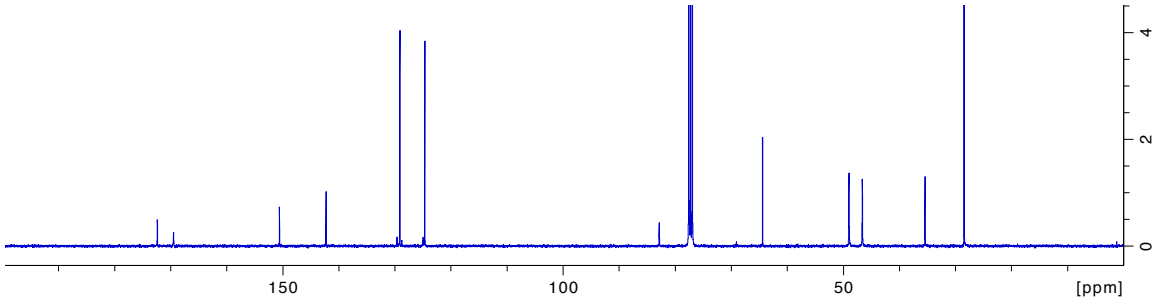


Data for L-14n. Yield 56%. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.39 (d, *J* = 8.7 Hz, 2H), 794 (d, *J* = 8.7 Hz, 2H), 6.04 (br, 1H), 5.45 (br, 1H), 3.62 (t, *J* = 7.2 Hz, 1H), 3.09 (m, 4H), 2.88 (m, 2H), 2.69 (m, 2H), 2.62 (dd, *J* = 15.6, 7.0 Hz, 1H), 2.43 (dd, *J* = 15.6, 7.3 Hz, 1H), 1.47 (s, 9H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 172.4, 169.5, 150.6, 142.3, 129.1, 124.6, 82.8, 64.4, 49.0, 46.6, 35.4, 28.4; HRMS (ESI) m/z calcd for $C_{18}H_{27}N_4O_7S^+$ 443.1595; found 443.1591 $(M+H)^+$

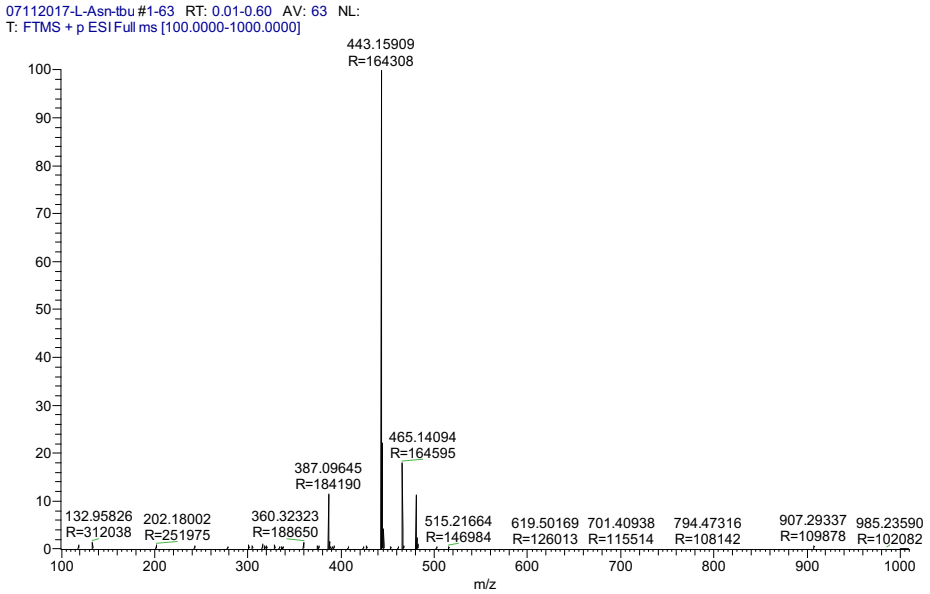
^{1}H NMR spectrum:



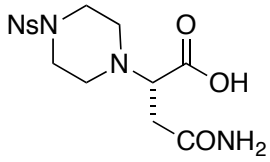
^{13}C NMR spectrum:



HRMS

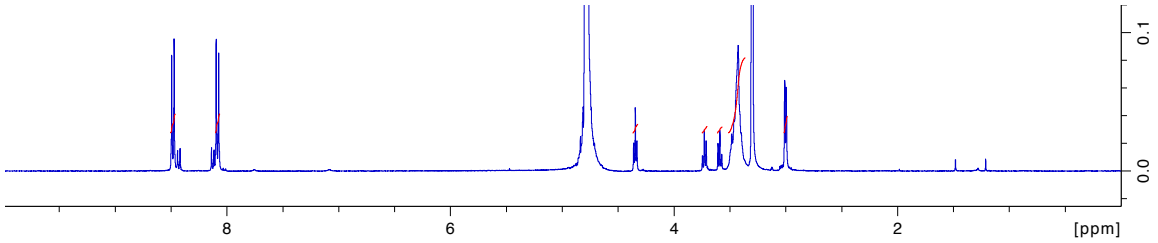


(*S*)-4-Amino-2-(4-((4-nitrophenyl)sulfonyl)piperazin-1-yl)-4-oxobutanoic acid (L-15n)

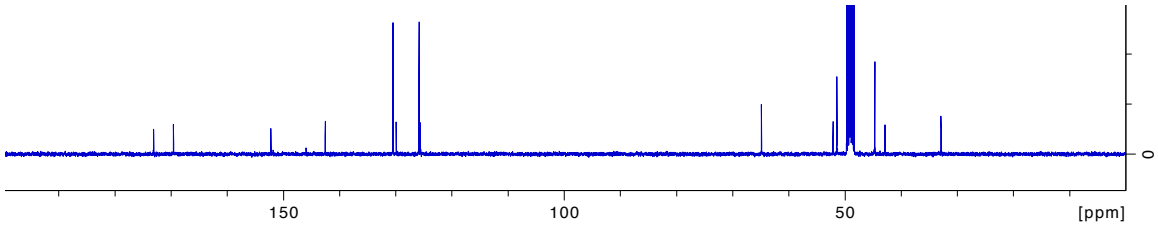


Data for L-15n. Yield 99%. ^{1}H NMR (400 MHz, MeOD-d_4) δ 8.48 (d, *J* = 8.9 Hz, 2H), 8.08 (d, *J* = 8.9 Hz, 2H), 4.35 (t, *J* = 5.7 Hz, 1H), 3.73 (t, *J* = 6.6 Hz, 1H), 3.59 (t, *J* = 6.6 Hz, 1H), 3.51-3.36 (m, 8H), 3.00 (d, *J* = 5.6 Hz, 2H) MeOD-d_4; ^{13}C NMR (100 MHz, MeOD-d_4) δ 173.1, 169.5, 152.2, 142.5, 130.5, 125.8, 64.9, 51.4, 44.7, 32.9; HRMS (ESI) m/z calcd for $C_{14}H_{19}N_4O_7S^+$ 387.0969; found 387.0966 $(M+H)^+$

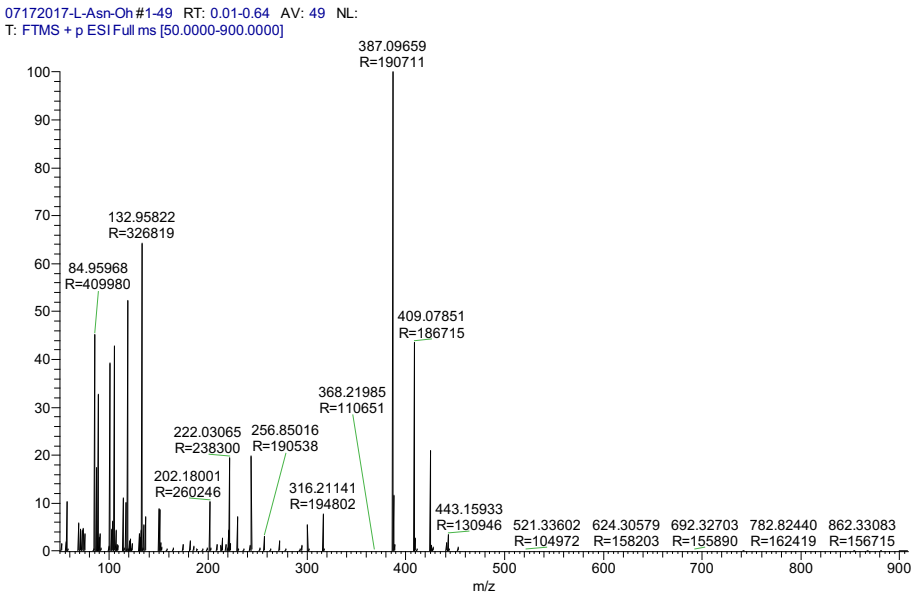
^{1}H NMR spectrum



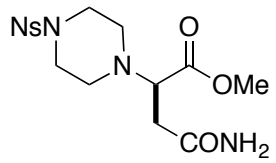
^{13}C NMR spectrum



HRMS

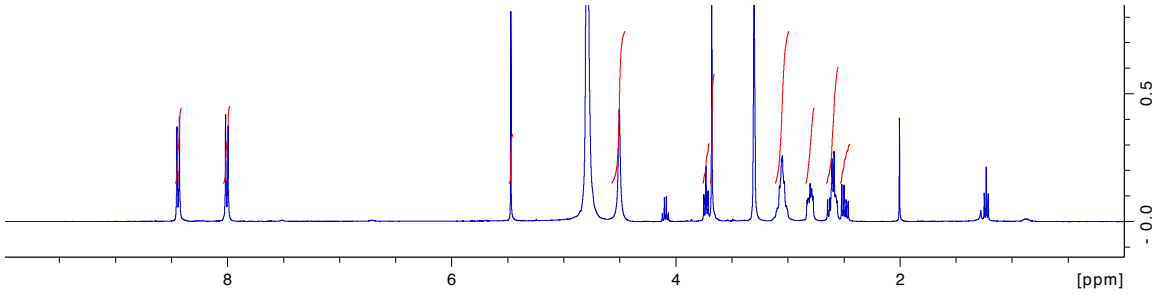


Methyl (*R*)-4-amino-2-(4-((4-nitrophenyl)sulfonyl)piperazin-1-yl)-4-oxobutanoate (D-14n)

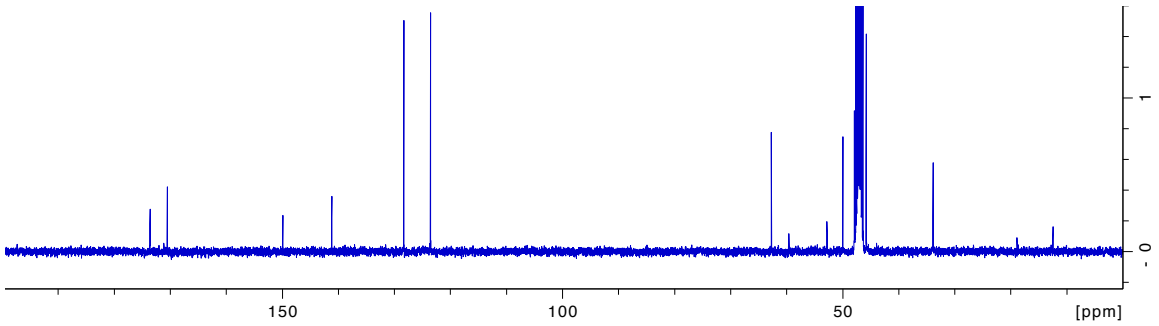


Data for D-14n. Yield 26%. ^{1}H NMR (400 MHz, MeOD-d_4) δ 8.43 (d, *J* = 8.9 Hz, 2H), 8.00 (d, *J* = 8.8 Hz, 2H), 3.73 (t, *J* = 7.6 Hz, 1H), 3.68 (s, 3H), 3.05 (m, 4H), 2.81 (m, 2H), 2.61 (m, 3H), 2.49 (dd, *J* = 15.2, 7.8 Hz, 1H); ^{13}C NMR (100 MHz, MeOD-d_4) δ 173.6, 170.5, 149.9, 141.1, 128.3, 123.5, 62.7, 49.9, 47.9, 45.8, 33.8; HRMS (ESI) m/z calcd for $C_{15}H_{21}N_4O_7S^+$ 401.1125; found 401.1123 $(M+H)^+$

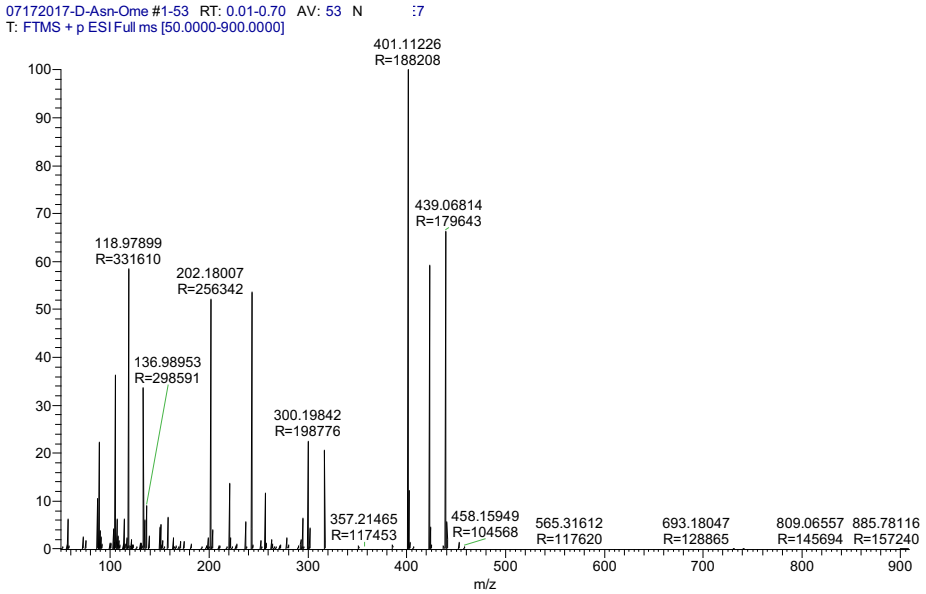
^{1}H NMR spectrum



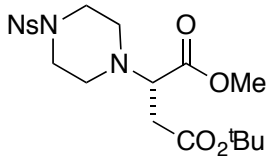
^{13}C NMR spectrum



HRMS

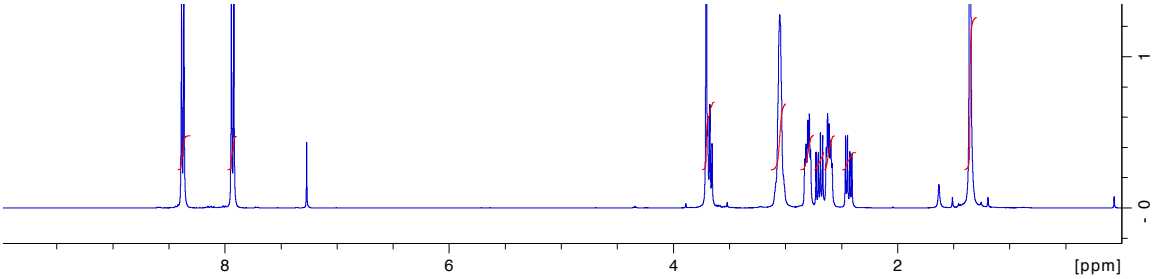


4-(*tert*-Butyl) 1-methyl (*S*)-2-(4-((4-nitrophenyl)sulfonyl)piperazin-1-yl)succinate (L-14d')

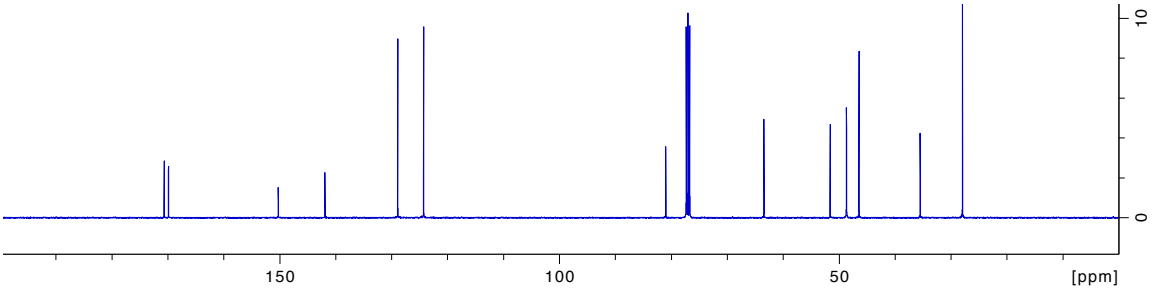


Data for L-14d'. Yield 65%. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.38 (d, *J* = 8.8 Hz, 2H), 7.93 (d, *J* = 8.8 Hz, 2H), 3.71 (s, 3H), 3.67 (m, 1H), 3.04 (m, 4H), 2.80 (m, 2H), 2.69 (dd, *J* = 15.7, 8.2 Hz, 1H), 2.62 (m, 2H), 2.43 (dd, *J* = 15.7, 7.0 Hz, 1H), 1.35 (s, 9H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 170.6, 169.9, 150.2, 141.9, 128.9, 124.3, 81.0, 63.4, 51.6, 48.7, 46.4, 35.5, 27.9; HRMS (ESI) m/z calcd for $C_{19}H_{28}N_3O_8S^+$ 458.1592; found 458.1587 $(M+H)^+$

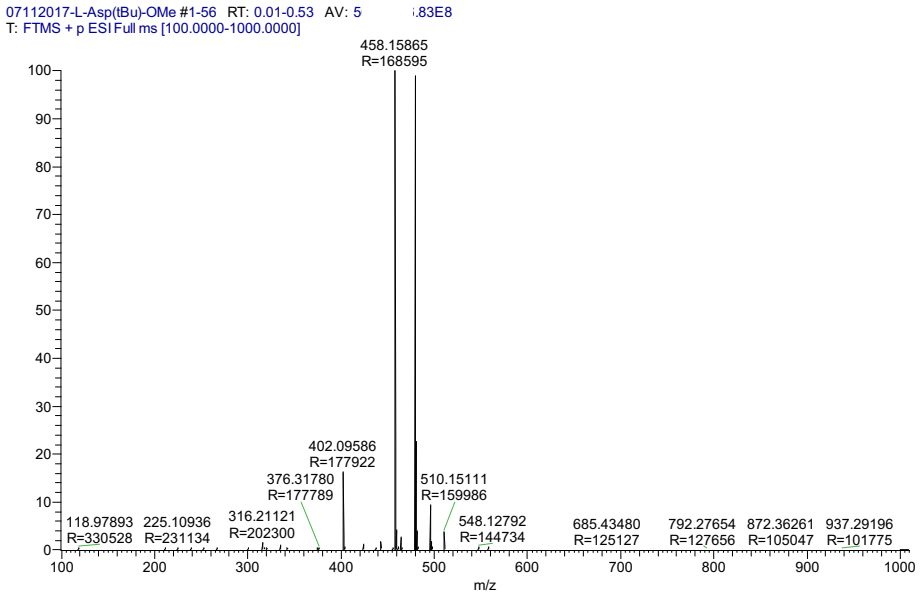
^{1}H NMR spectrum



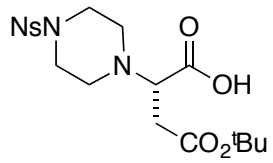
^{13}C NMR spectrum



HRMS

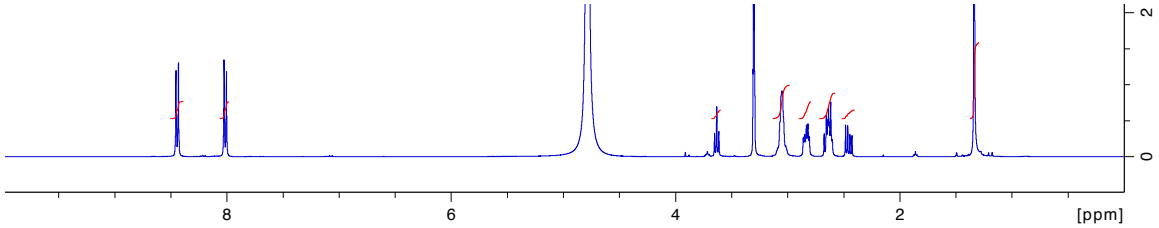


(*S*)-4-(*tert*-Butoxy)-2-(4-((4-nitrophenyl)sulfonyl)piperazin-1-yl)-4-oxobutanoic acid (L-15d')

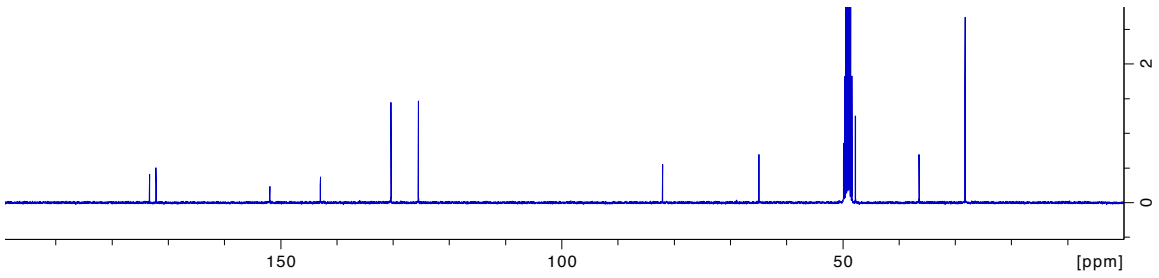


Data for L-15d'. Yield 81%. ^{1}H NMR (400 MHz, MeOD-d_4) δ 8.44 (d, *J* = 8.8 Hz, 2H), 8.01 (d, *J* = 8.9 Hz, 2H), 3.63 (t, *J* = 7.7 Hz, 1H), 3.05 (m, 4H), 2.84 (m, 2H), 2.64 (m, 3H), 2.45 (dd, *J* = 15.5, 7.5 Hz, 1H), 1.34 (s, 9H); ^{13}C NMR (100 MHz, MeOD-d_4) δ 173.1, 172.0, 151.9, 142.9, 130.4, 125.5, 82.0, 64.9, 49.9, 47.7, 36.3, 28.2; HRMS (ESI) m/z calcd for $C_{18}H_{25}N_3O_8S$ 444.1435; found 444.1430 $(M+H)^+$

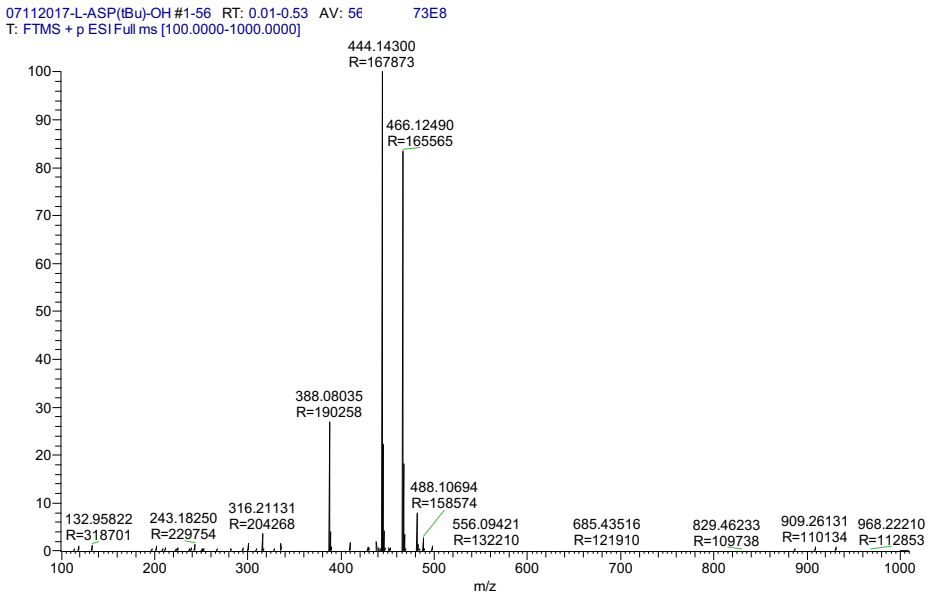
^{1}H NMR spectrum:



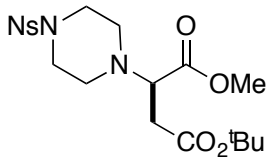
^{13}C NMR spectrum



HRMS

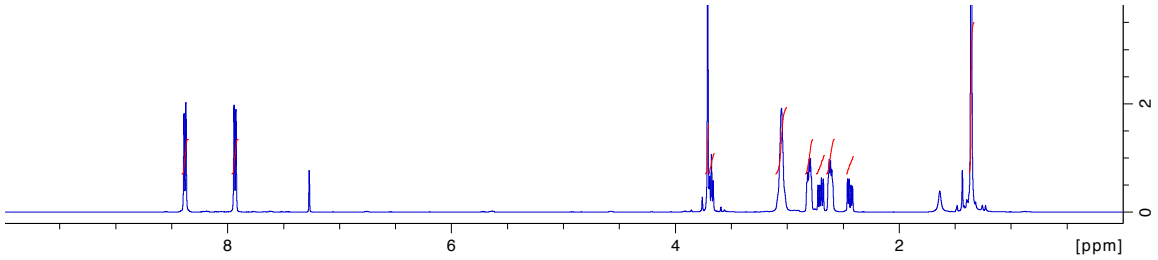


4-(*tert*-Butyl) 1-methyl (*R*)-2-(4-((4-nitrophenyl)sulfonyl)piperazin-1-yl)succinate (D-14d')

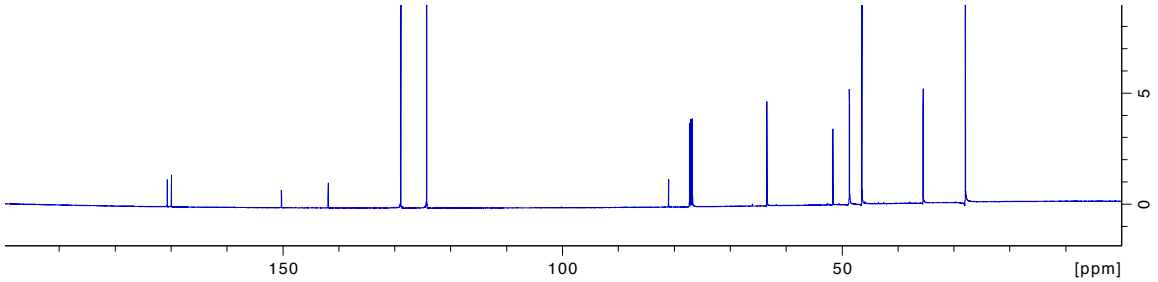


Data for D-14d'. Yield 42%. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.38 (d, *J* = 8.0 Hz, 2H), 7.93 (d, *J* = 8.0 Hz, 2H), 3.71 (s, 3H), 3.68 (t, *J* = 7.7 Hz, 1H), 3.05 (m, 4H), 2.81 (m, 2H), 2.70 (dd, *J* = 15.7, 8.1 Hz, 1H), 2.61 (m, 2H), 2.44 (dd, *J* = 15.7, 6.9 Hz, 1H), 1.36 (s, 9H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 170.6, 169.9, 150.2, 141.9, 128.9, 124.3, 81.0, 63.4, 51.6, 48.7, 46.4, 35.5, 27.9; HRMS (ESI) m/z calcd for $C_{19}H_{28}N_3O_8S^+$ 458.1592; found 458.1593 $(M+H)^+$

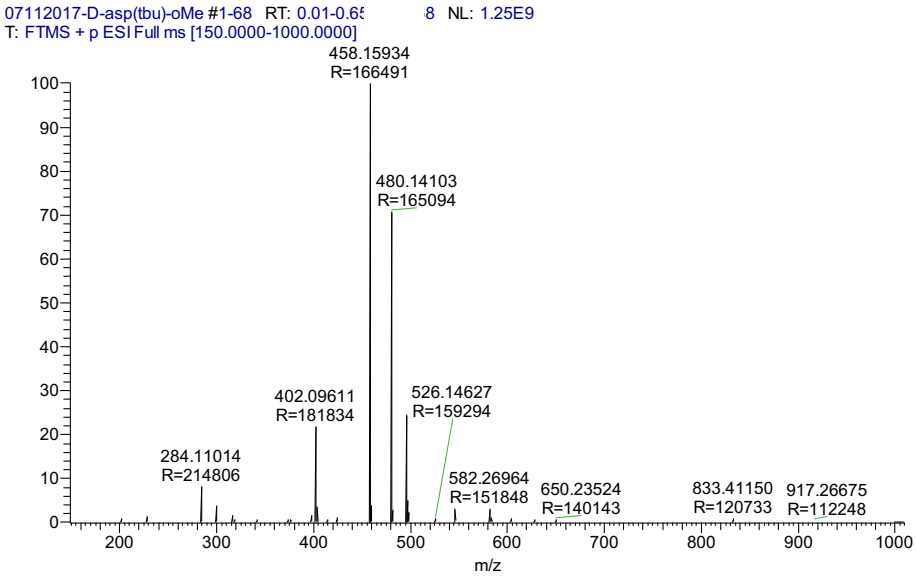
^{1}H NMR spectrum:



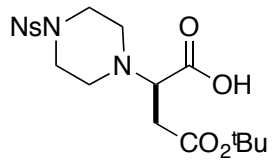
^{13}C NMR spectrum:



HRMS

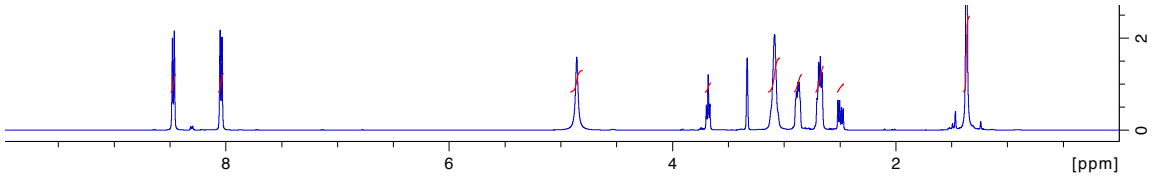


(*R*)-4-(*tert*-Butoxy)-2-(4-((4-nitrophenyl)sulfonyl)piperazin-1-yl)-4-oxobutanoic acid (D-15d')

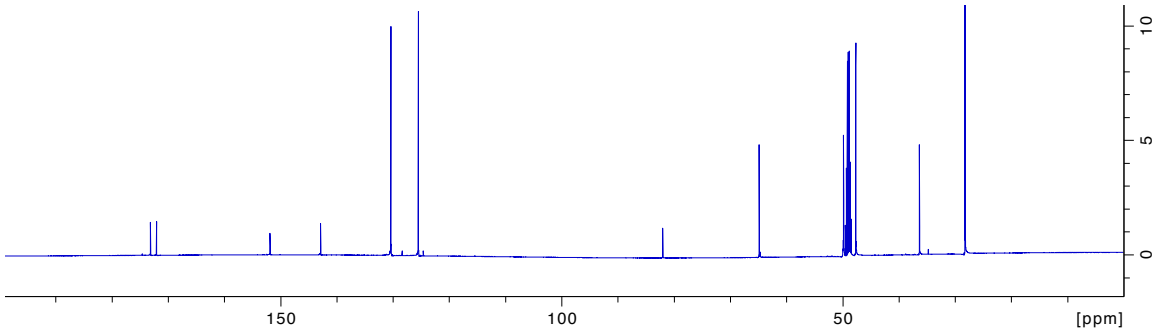


Data for D-15d'. Yield 66%. ^{1}H NMR (400 MHz, MeOD-d_4) δ 8.47 (d, *J* = 8.5 Hz, 2H), 8.04 (d, *J* = 8.5 Hz, 2H), 3.68 (t, *J* = 7.6 Hz, 1H), 3.08 (m, 4H), 2.88 (m, 2H), 2.68 (m, 3H), 2.49 (dd, *J* = 15.6, 7.4 Hz, 1H), 1.37 (s, 9H); ^{13}C NMR (100 MHz, MeOD-d_4) δ 173.1, 172.0, 151.9, 142.9, 130.4, 125.5, 82.0, 64.9, 49.9, 47.7, 36.3, 28.2; HRMS (ESI) m/z calcd for $C_{18}H_{26}N_3O_8S^+$ 444.1435; found 444.1434 $(M+H)^+$

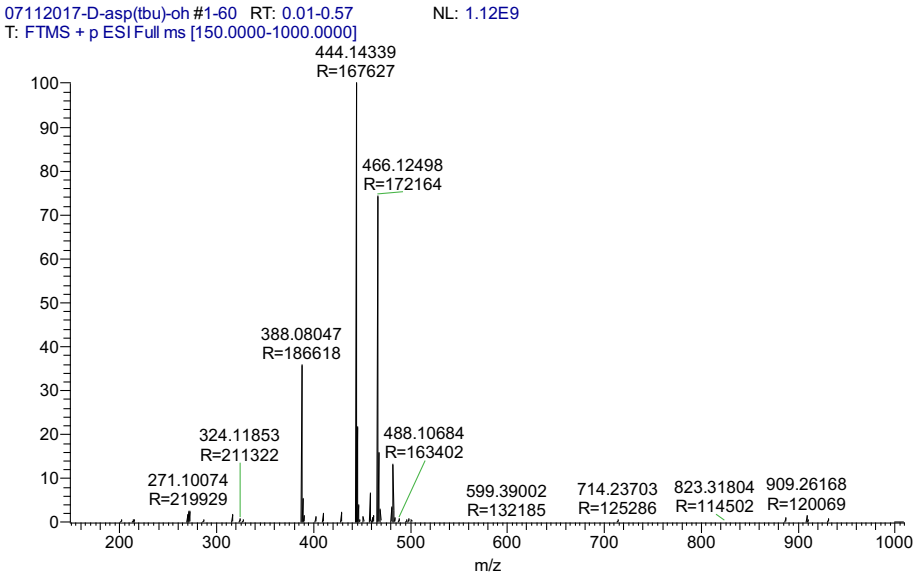
^{1}H NMR spectrum:



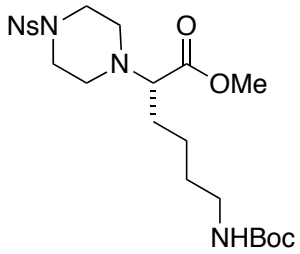
^{13}C NMR spectrum:



HRMS

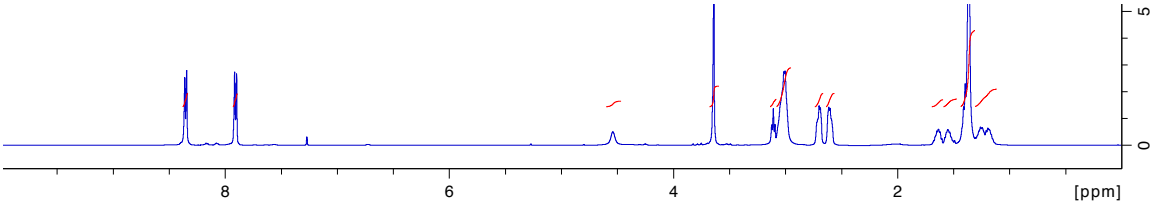


Methyl (*S*)-6-((*tert*-butoxycarbonyl)amino)-2-(4-((4-nitrophenyl)sulfonyl)piperazin-1-yl)hexanoate (L-14k')

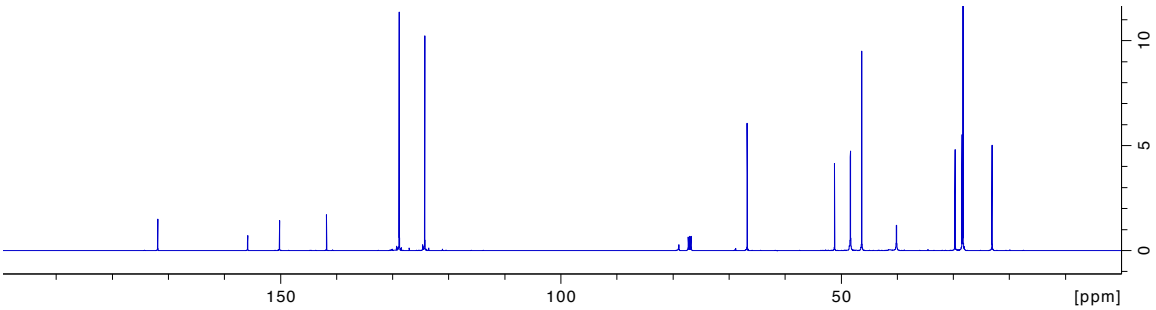


Data for L-14k'. Yield 58%. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.35 (d, *J* = 8.6 Hz, 2H), 7.91 (d, *J* = 8.6 Hz, 2H), 4.54 (br, 1H), 3.64 (s, 3H), 3.11 (t, *J* = 7.3 Hz, 1H), 3.01 (m, 6H), 2.71 (m, 2H), 2.60 (m, 2H), 1.63 (m, 1H), 1.55 (m, 1H), 1.43-1.31 (m, 11H), 1.31-1.11(m, 2H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 171.9, 155.8, 150.1, 141.8, 128.8, 124.6, 79.0, 66.7, 51.2, 48.3, 46.3, 40.1, 29.7, 28.4, 28.2, 23.1 HRMS (ESI) m/z calcd for $C_{22}H_{35}N_4O_8S^+$ 515.2170; found 515.2172 $(M+H)^+$

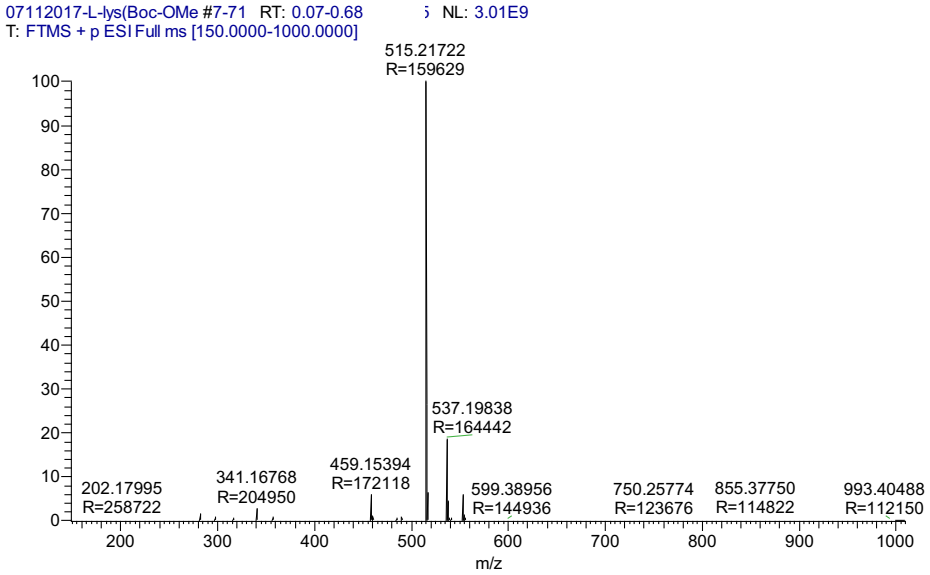
^{1}H NMR spectrum



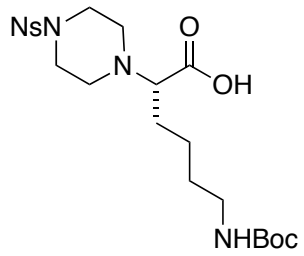
^{13}C NMR spectrum



HRMS

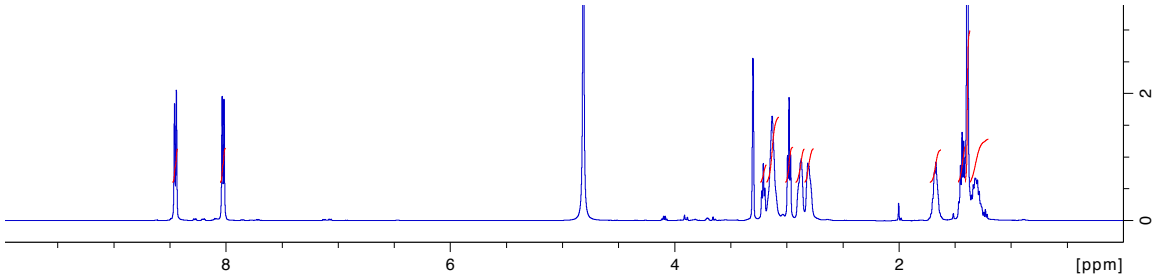


(*S*)-6-((*tert*-Butoxycarbonyl)amino)-2-(4-((4-nitrophenyl)sulfonyl)piperazin-1-yl)hexanoic acid (L-15k')

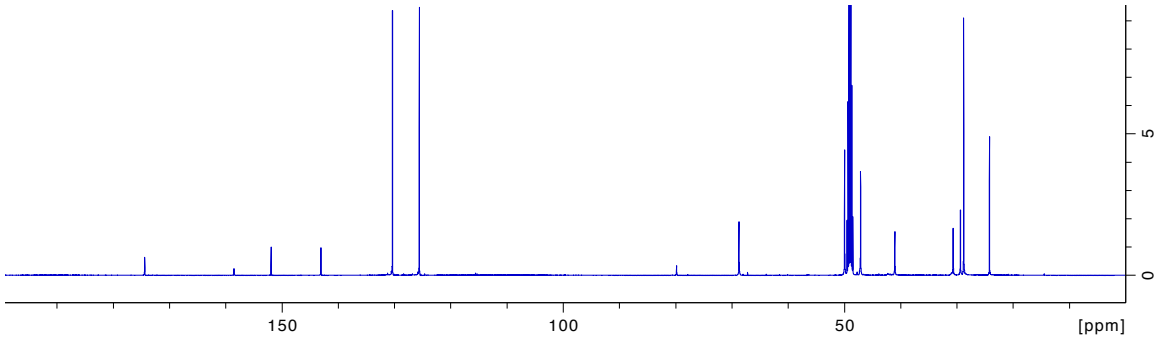


Data for L-15k'. Yield 100%. ^{1}H NMR (400 MHz, MeOD-d_4) δ 8.45 (d, *J* = 8.0 Hz, 2H), 8.02 (d, *J* = 8.0 Hz, 2H), 3.21 (t, *J* = 6.9 Hz, 1H), 3.15 (m, 4H), 2.98 (t, *J* = 6.7 Hz, 2H), 2.87 (m, 2H), 2.81 (m, 2H), 1.68 (m, 2H), 1.44 (m, 2H), 1.39 (s, 9H), 1.31 (m, 2H); ^{13}C NMR (100 MHz, MeOD-d_4) δ 174.4, 158.5, 151.9, 143.1, 130.4, 125.6, 79.8, 68.7, 49.9, 47.1, 41.0, 30.7, 29.4, 28.8, 24.2; HRMS (ESI) m/z calcd for $C_{21}H_{33}N_4O_8S^+$ 501.2014; found 501.2012 $(M+H)^+$

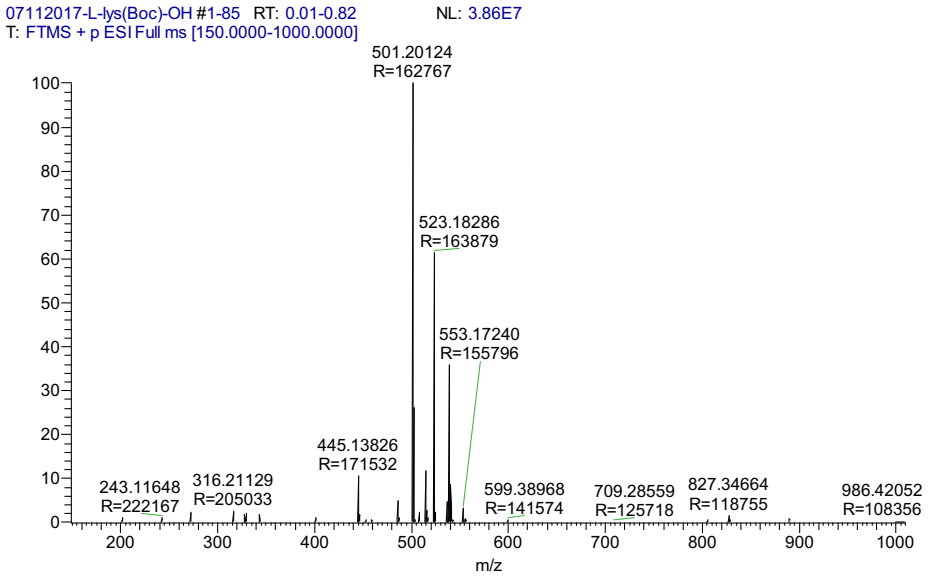
^{1}H NMR spectrum



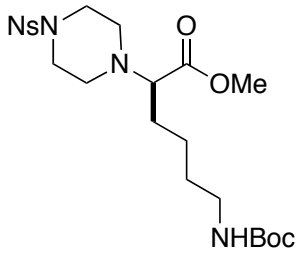
^{13}C NMR spectrum



HRMS

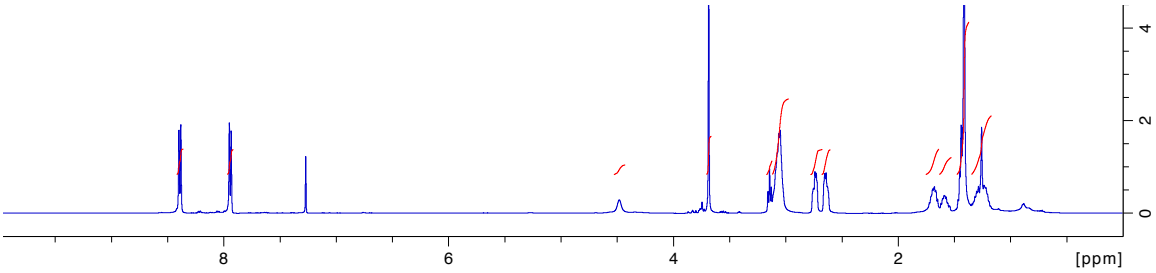


Methyl (*R*)-6-((*tert*-butoxycarbonyl)amino)-2-(4-((4-nitrophenyl)sulfonyl)piperazin-1-yl)hexanoate (D-14k')

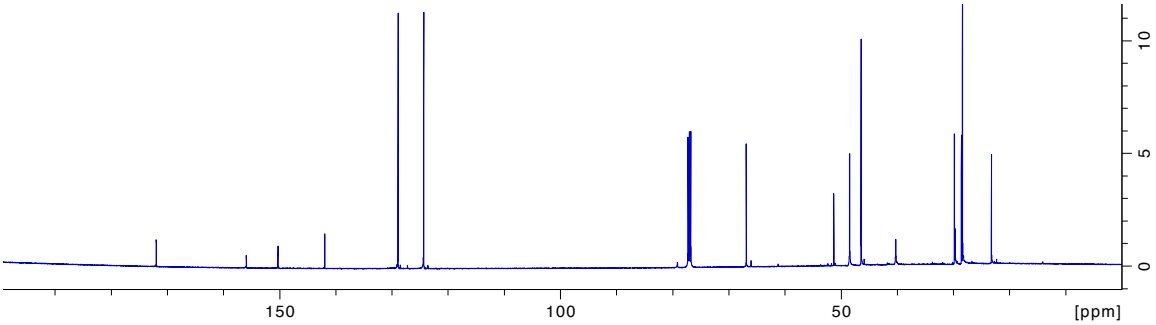


Data for D-14k'. Yield 60%. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.39 (d, *J* = 8.3 Hz, 2H), 7.94 (d, *J* = 8.3 Hz, 2H), 4.48 (br, 1H), 3.69 (s, 3H), 3.14 (t, *J* = 7.4 Hz, 1H), 3.07 (m, 6H), 2.74 (m, 2H), 2.64 (m, 2H), 1.68 (m, 2H), 1.59 (m, 1H), 1.42 (m, 10 H), 1.28 (m, 2H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 172.0, 155.9, 150.3, 141.9, 128.9, 124.3, 79.2, 66.9, 51.3, 48.5, 46.4, 40.2, 29.8, 28.6, 28.4, 23.2; HRMS (ESI) m/z calcd for $C_{22}H_{35}N_4O_8S^+$ 515.2170; found 515.2168 $(M+H)^+$

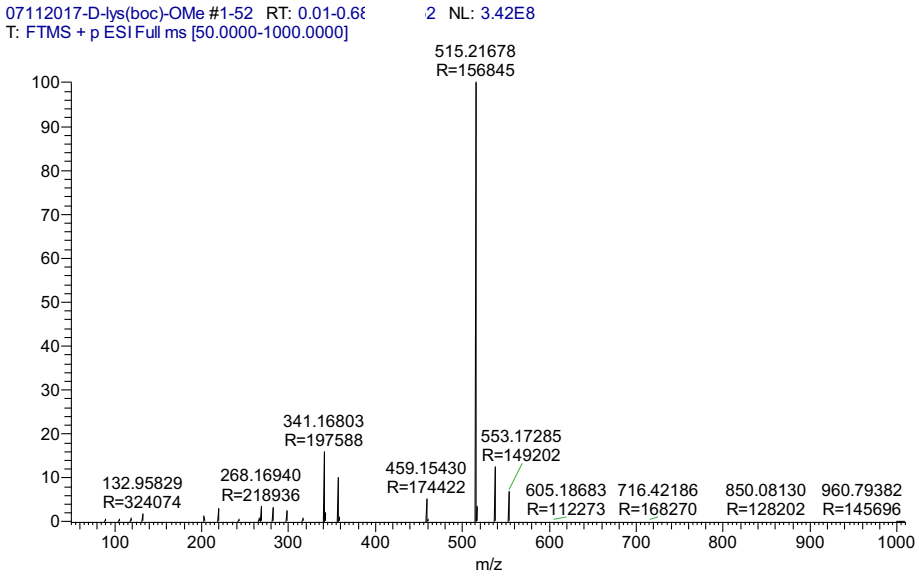
^{1}H NMR spectrum:



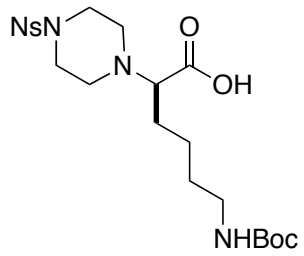
^{13}C NMR spectrum:



HRMS

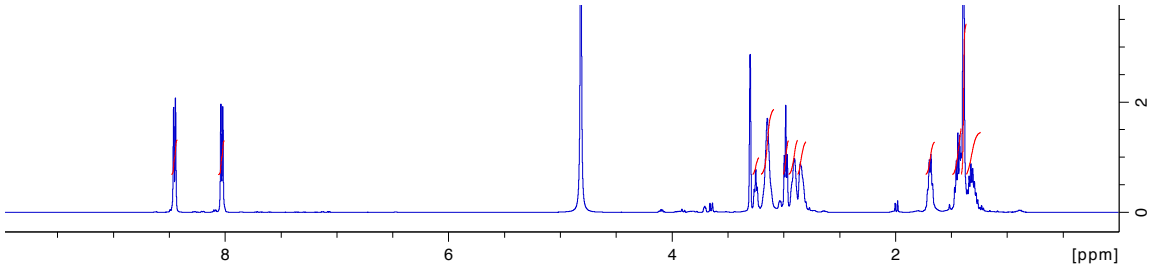


(*R*)-6-((*tert*-Butoxycarbonyl)amino)-2-(4-((4-nitrophenyl)sulfonyl)piperazin-1-yl)hexanoic acid (D-15k')

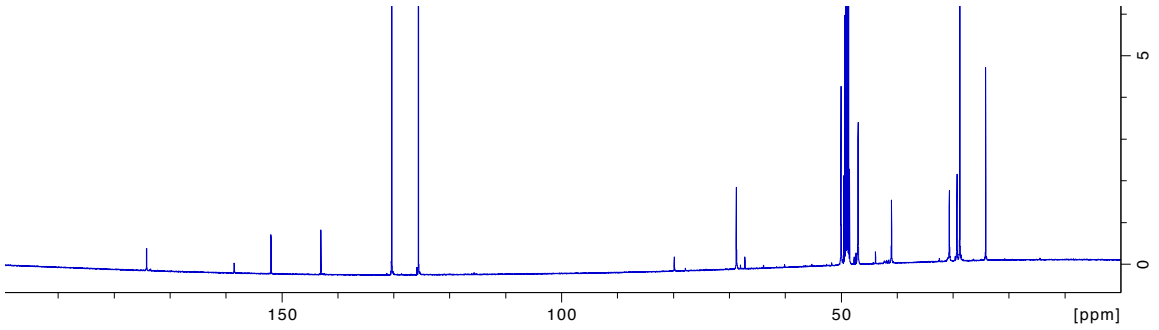


Data for D-15k'. Yield 70%. ^{1}H NMR (400 MHz, MeOD-d_4) δ 8.45 (d, *J* = 7.9 Hz, 2H), 8.03 (d, *J* = 7.9 Hz, 2H), 3.25 (t, *J* = 6.8 Hz, 1H), 3.15 (m, 4H), 2.98 (t, *J* = 6.7 Hz, 2H), 2.91 (m, 2H), 2.84 (m, 2H), 1,69 (m, 2H), 1.49-1.41 (m, 3H), 1.39 (s, 9H), 1.31 (m, 2H); ^{13}C NMR (100 MHz, MeOD-d_4) δ 174.2, 158.8, 151.9, 143.0, 130.4, 125.6, 79.8, 68.7, 50.0, 47.0, 41.0, 30.6, 29.3, 28.8, 24.2; HRMS (ESI) m/z calcd for $C_{21}H_{33}N_4O_8S^+$ 501.2014; found 501.2014 $(M+H)^+$

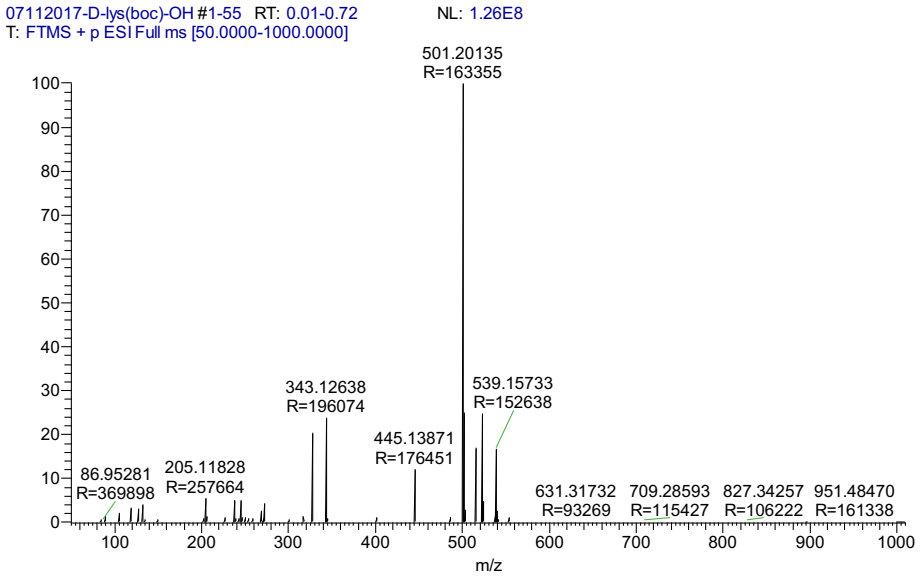
^{1}H NMR spectrum



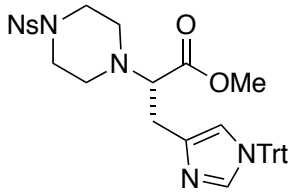
^{13}C NMR spectrum



HRMS

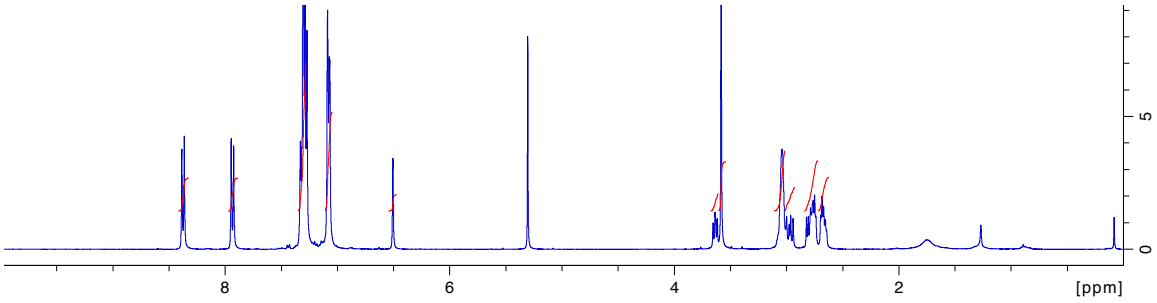


Methyl (*S*)-2-(4-((4-nitrophenyl)sulfonyl)piperazin-1-yl)-3-(1-trityl-1*H*-imidazol-4-yl)propanoate (L-14h')

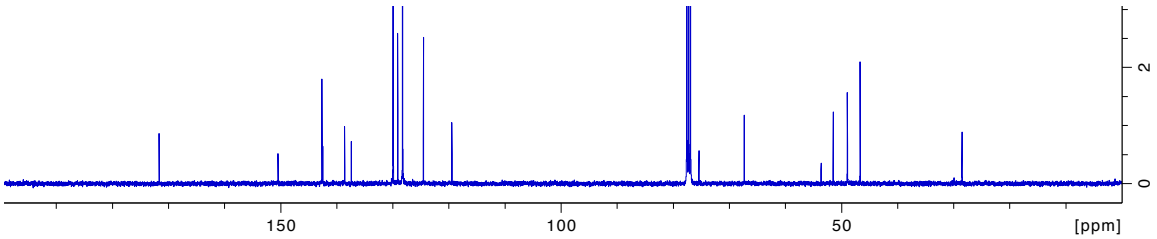


Data for L-14h'. Yield 39%. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.37 (d, *J* = 8.7 Hz, 2H), 7.93 (d, *J* = 8.7 Hz, 2H), 7.35-7.24 (m, 10H), 7.08 (m, 6H), 6.51 (s, 1H), 3.63 (dd, *J* = 8.6, 6.5 Hz, 1H), 3.58 (s, 3H), 3.04 (m, 4H), 2.97 (dd, *J* = 14.3, 8.9 Hz, 1H), 2.84-2.72 (m, 3H), 2.68 (m, 2H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 171.7, 150.4, 142.6, 142.5, 138.6, 137.4, 129.9, 129.1, 128.2, 124.5, 119.4, 75.4, 67.3, 53.6, 51.5, 48.9, 46.7, 28.5; HRMS (ESI) m/z calcd for $C_{36}H_{36}N_5O_6S^+$ 666.2381; found 666.2373 $(M+H)^+$

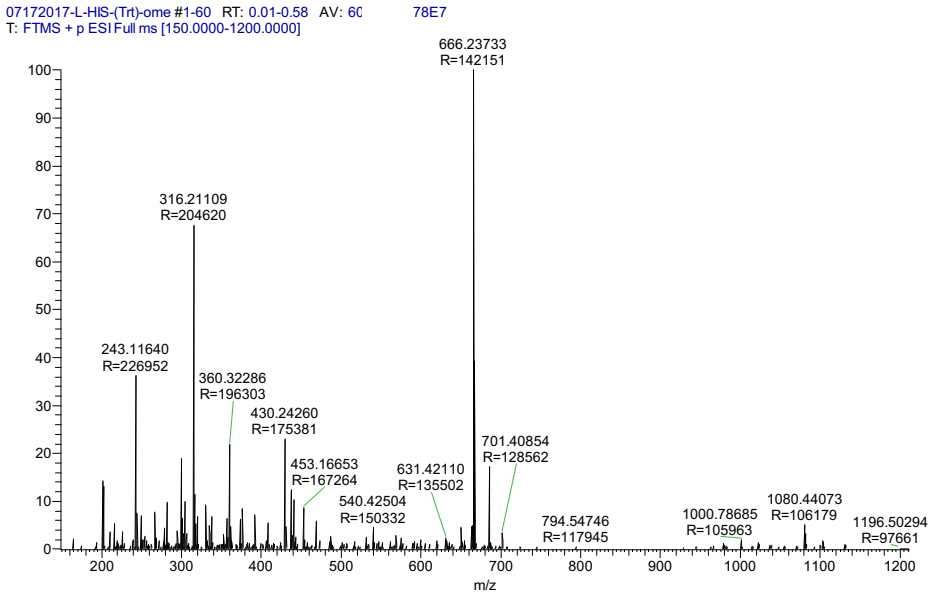
^{1}H NMR spectrum



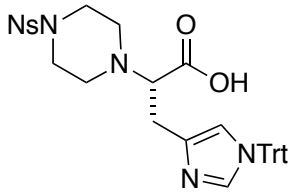
^{13}C NMR spectrum



HRMS

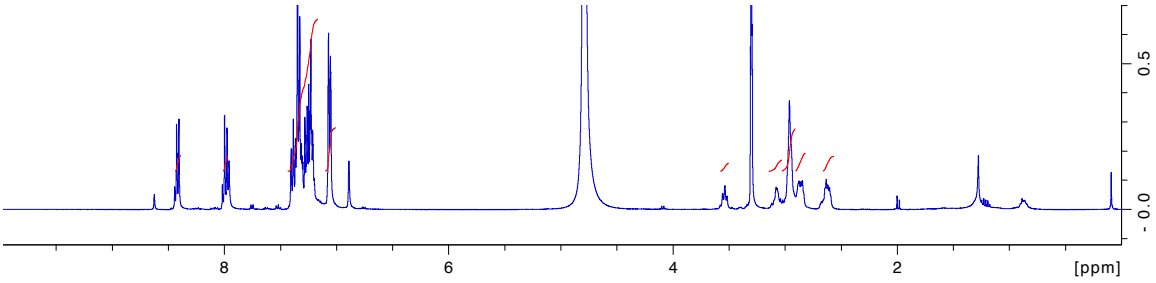


(*S*)-2-(4-((4-Nitrophenyl)sulfonyl)piperazin-1-yl)-3-(1-trityl-1*H*-imidazol-4-yl)propanoic acid (L-15h')

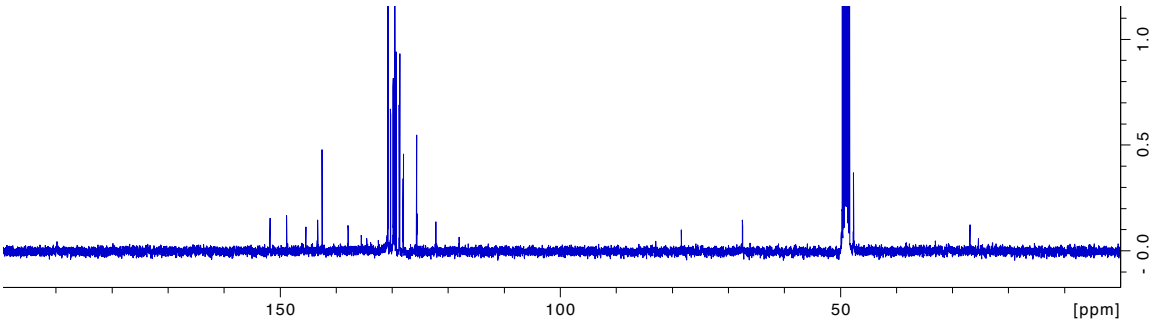


Data for L-15h'. Yield 100%. ^{1}H NMR (400 MHz, MeOD-d_4) δ 8.41 (d, *J* = 8.8 Hz, 2H), 7.99 (d, *J* = 8.8 Hz, 2H), 7.43-7.17 (m, 11H), 7.06 (m, 5H), 6.89 (m, 1H), 3.53 (t, *J* = 7.7 Hz, 1H), 3.08 (m, 1H), 2.96 (m, 5H), 2.87 (m, 2H), 2.63 (m, 2H); ^{13}C NMR (100 MHz, MeOD-d_4) δ 151.9, 148.9, 145.5, 143.3, 142.5, 137.8, 130.7, 130.3, 129.5, 129.3, 125.6, 122.2, 78.4, 67.4, 47.8, 26.9; HRMS (ESI) m/z calcd for $C_{35}H_{34}N_5O_6S^+$ 652.2224; found 652.2224 $(M+H)^+$

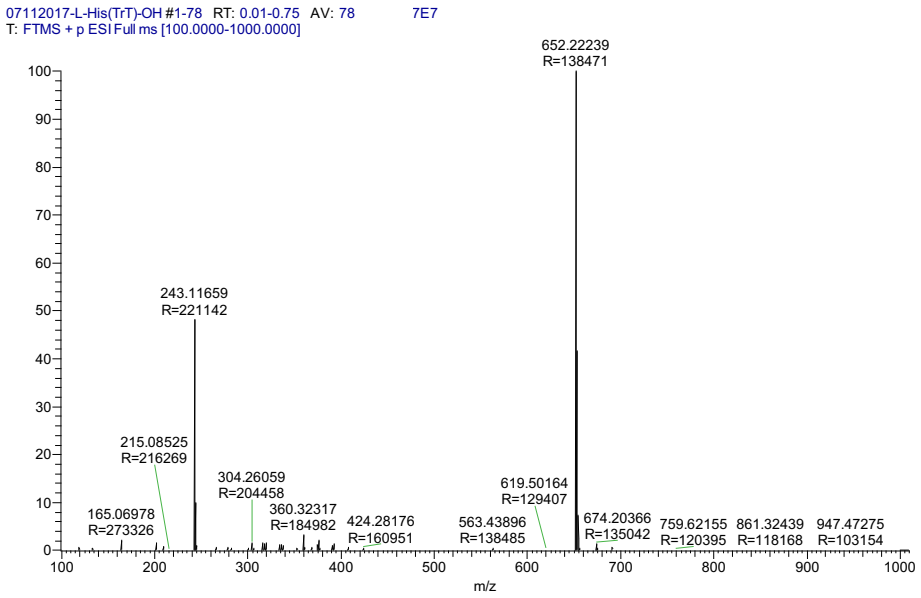
^{1}H NMR spectrum



^{13}C NMR spectrum



HRMS



Solid-phase Syntheses and Characterizations of Compounds ***12***

Loading pip-acids on resin: 2-Cl-Trt resin (200 mg, 1.2 mequiv/g) in fritted syringe was swollen in CH_2Cl_2 for 1 h. Pip-acid (1.2 eq) and $^{i}Pr_2NEt$ (3 eq) in 3 mL of CH_2Cl_2/DMF (1:1) were added into the syringe, and beads were subjected to the microwave irradiation (100 W, 50 ºC, 30 min). Loading process was repeated with a fresh solution. Beads were washed thrice with CH_2Cl_2 (4 mL x 3), then the remaining active beads were blocked with CH_2Cl_2 / MeOH / $^{i}Pr_2NEt$ (17:2:1) cocktail for 1 h. Beads were washed thoroughly with CH_2Cl_2 (4 mL x 5 mL), MeOH (4 mL x 3), CH_2Cl_2 (4 mL x 3) and DMF (4 mL x 3).

Nosyl deprotection: Beads containing nosyl-protected pip-acids were treated with the solution of mercaptoethanol (5 eq) and DBU (5 eq) in DMF for 15 minutes at 25 ºC. The

orange solution was drained, and the fresh deprotection cocktail was added to the beads. After the second treatment, beads were washed thoroughly with DMF (4 mL x 5). The completeness of reaction was monitored by chloranil test.

Amide coupling: Resin containing free-amine functional group was added with the solution of Fmoc-protected amino acids (4 eq), $^{i}Pr_2NEt$ (8 eq) and 0.25 M HATU/HOAt solution in DMF. Beads were subjected to microwave irradiation (100 W, 50 ºC, 15 min). Then, beads were washed with 4 mL x 5 of DMF. The completeness of amide coupling can be monitored by Kaiser test (for a primary amine) or chloranil test (for a secondary amine).

Fmoc deprotection: Beads were shaken in 2 mL of 20% piperidine/DMF for 1 minute at 25 ºC. Then, the solution was drained and the fresh deprotection solution was added. Beads were heated under microwave irradiation (160 W, 75 ºC, 3 min) to deprotect Fmoc group. Solution was drained, and washed thoroughly with DMF (4 mL x 5), CH_2Cl_2 (4 mL x 3), MeOH (2 mL x 3) and DMF (4 mL x 3). Positive blue color of Kaiser test was monitored for completeness of reaction.

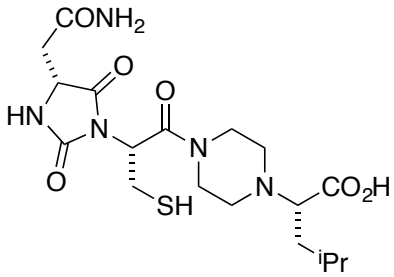
Hydantoin cyclization: After Fmoc deprotection at the R^1 position, beads were treated with the solution of *p*-nitrophenyl chloroformate (2 eq) and $^{i}Pr_2NEt$ (4 eq) in CH_2Cl_2 (~0.05 M), and were shaken for 30 minutes at 25 ºC. The solution was drained, and beads were treated again with a fresh coupling reagent. After that, beads were washed with CH_2Cl_2 (4 mL x 5) and NMP (4mL x 3). Solution of 5% $^{i}Pr_2NEt$ in NMP was added , and the beads were shaken for 1 h. The yellow solution was discarded, and the process was repeated with 5% $^{i}Pr_2NEt$/NMP for 4 more times. At the end of removing

byproducts, beads were washed with NMP (4 mL x 5), CH_2Cl_2 (4 mL x 5), MeOH (4 mL x 5) and dried in vacuum.

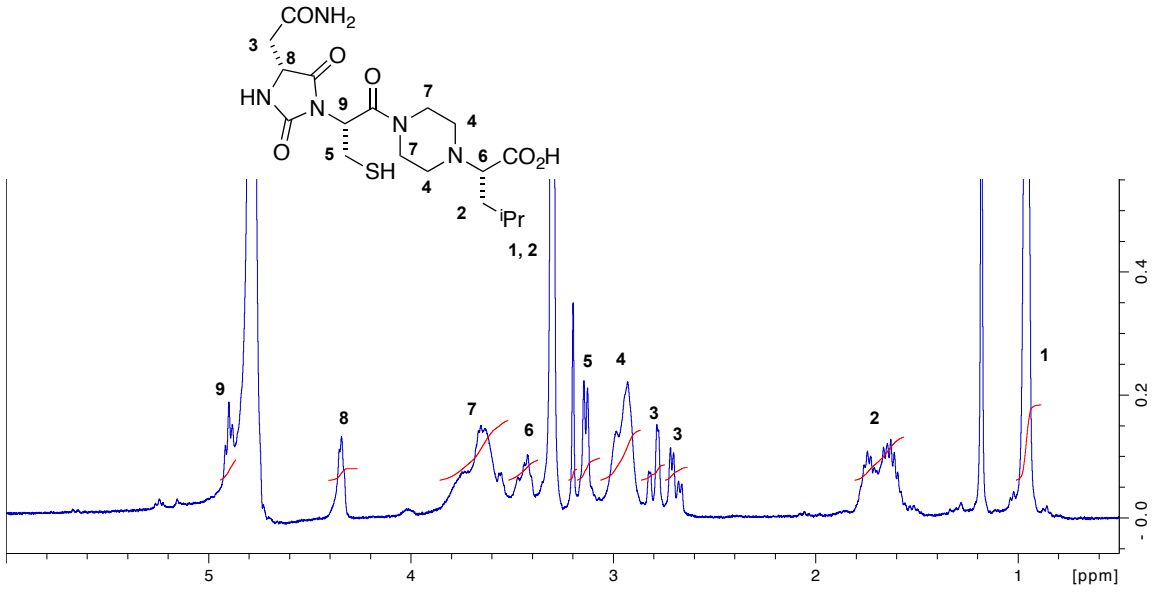
Cleavage from solid support: Products were cleaved from the beads with 20% HFIP/CH_2Cl_2 solution for 24 h. Crude material was dried under N_2 stream, and purified by preparative reverse-phase HPLC (50-95% MeCN/water containing 0.05% TFA). The purified products were lyophilized to obtain protected-form of compounds **12** as white powders.

Side-chain deprotection: Purified lyophilized compounds were stirred in TFA/Et_3SiH/H_2O (95: 2.5: 2.5 v/v) cocktail for 4 h at room temperature to remove acid-labile protecting group. Crude materials were dried under vacuum, and then precipitated with cold *tert*-butyl methyl ether. Solid compounds were collected, washed several times with ether and the final products were lyophilized to obtain products as white powders.

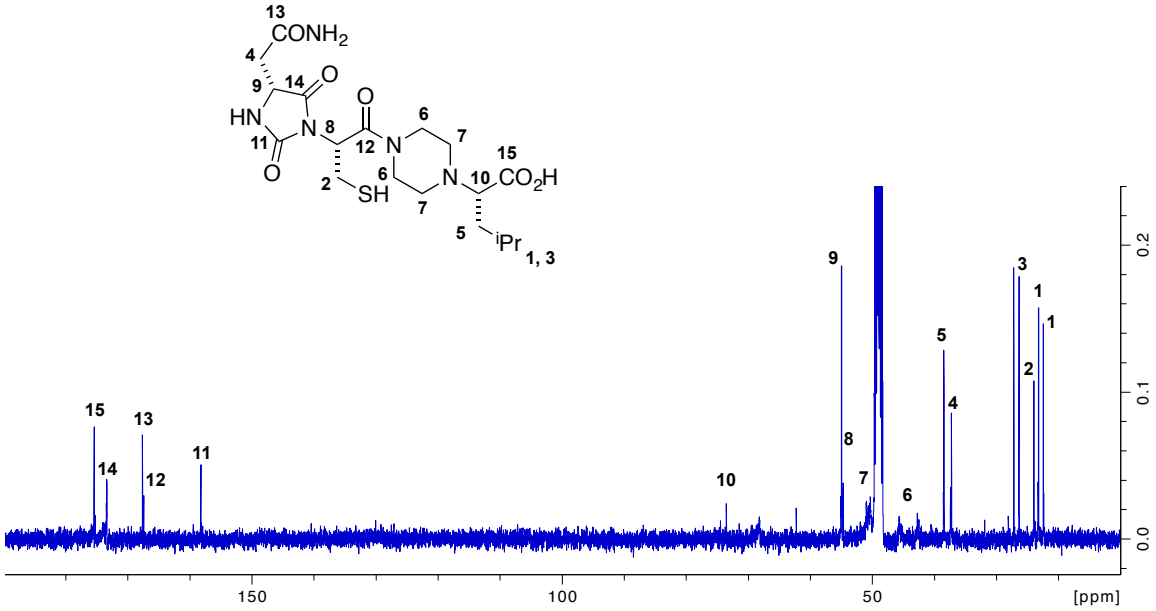
(*S*)-2-(4-((*R*)-2-((*R*)-4-(2-Amino-2-oxoethyl)-2,5-dioxoimidazolidin-1-yl)-3-mercaptopropanoyl)piperazin-1-yl)-4-methylpentanoic acid (DLL-12ncl)



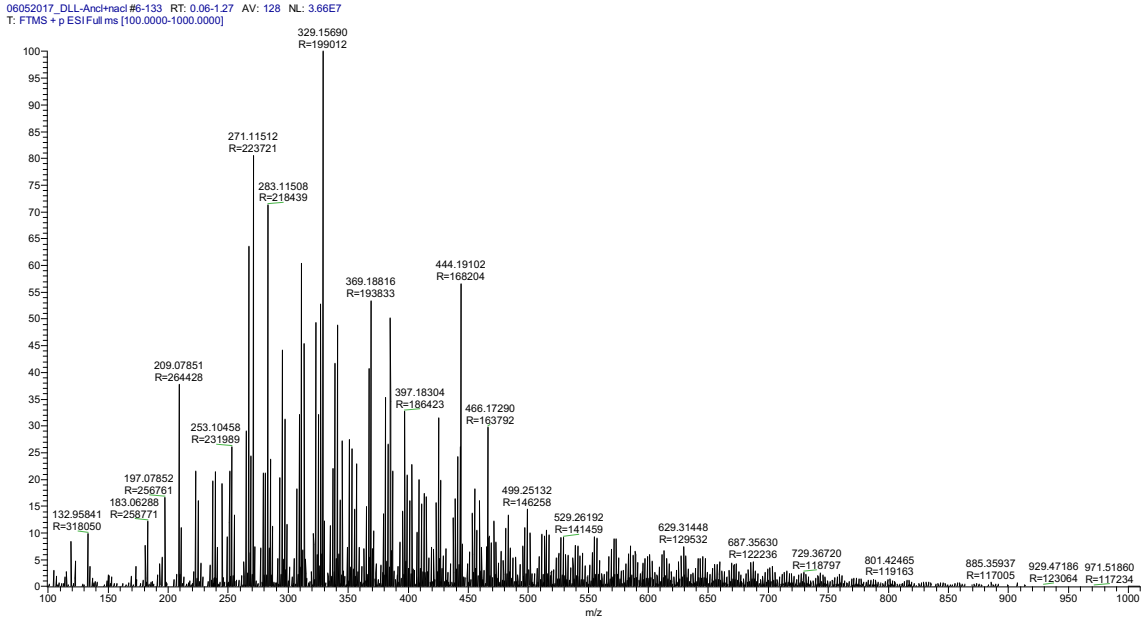
Data for DLL-12ncl. ^{1}H NMR (400 MHz, MeOD-d_4) δ 4.91 (m, 1H), 4.34 (m, 1H), 3.85-3.50 (m, 4H), 3.50-3.36 (m, 1H), 3.13 (d, *J* = 7.5 Hz, 2H), 3.02-2.85 (m, 4H), 2.80 (dd, *J* = 16.1, 4.2 Hz, 1H), 2.68 (dd, *J* = 16.1, 6.5 Hz, 1H), 1.81-1.49 (m, 3H), 0.96 (m, 6H);



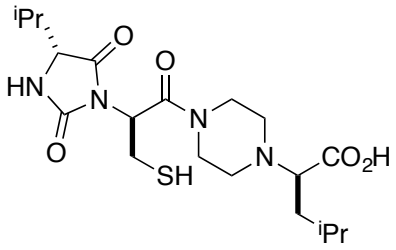
^{13}C NMR (100 MHz, MeOD-d_4) δ 175.4, 173.4, 167.6, 167.4, 158.2, 73.5, 55.0, 54.7, 50.9, 45.7, 42.7, 38.5, 37.3, 27.2, 26.4, 24.0, 23.2, 22.4



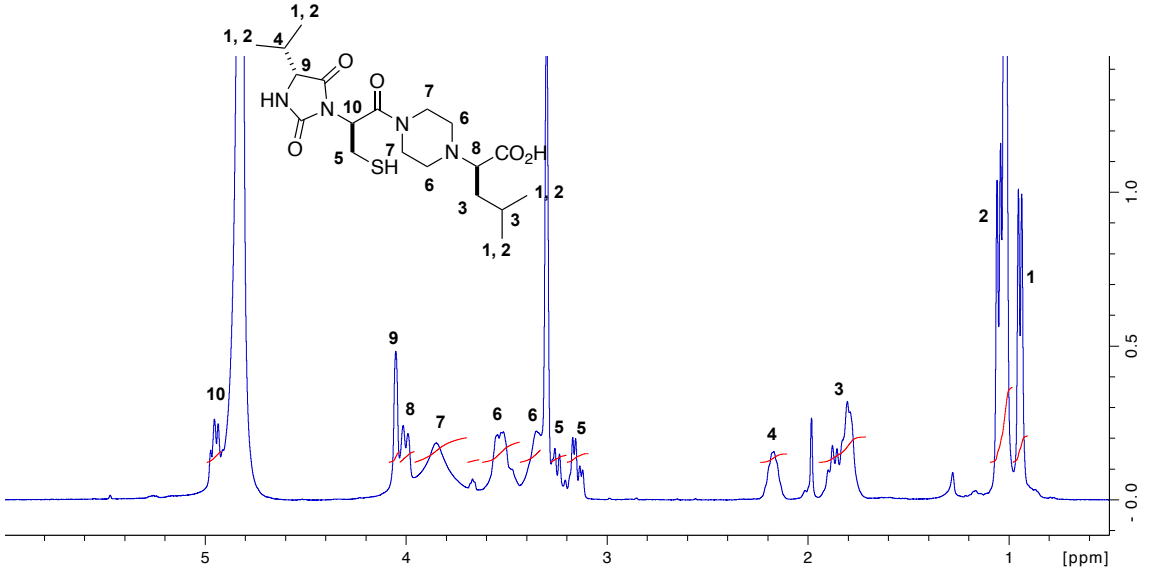
HRMS (ESI) m/z calcd for $C_{18}H_{30}N_5O_6S^+$ 444.1911; found 444.1910 $(M+H)^+$



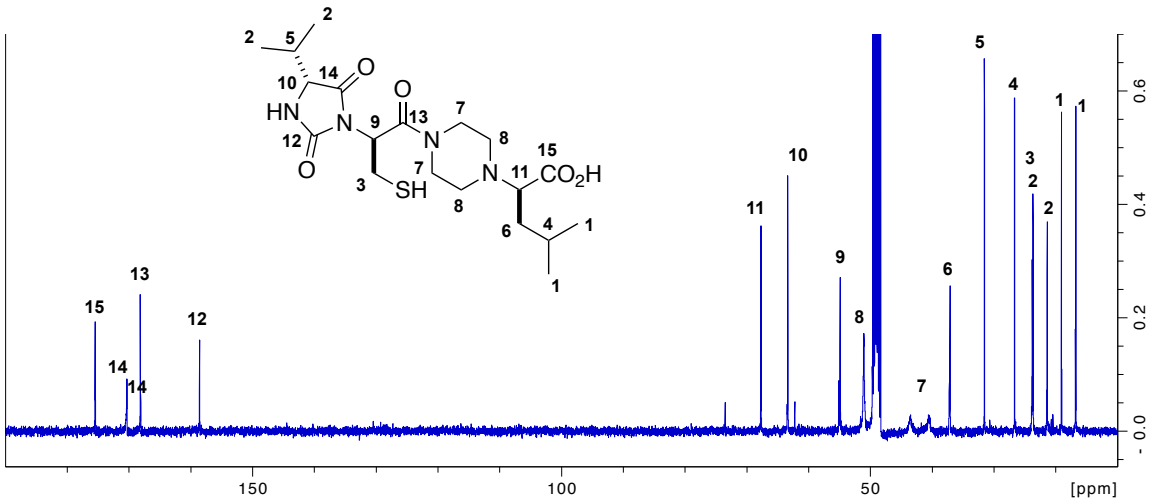
(*R*)-2-(4-((*S*)-2-((*R*)-4-Isopropyl-2,5-dioxoimidazolidin-1-yl)-3-mercaptopropanoyl)piperazin-1-yl)-4-methylpentanoic acid (DDD-12vcl)



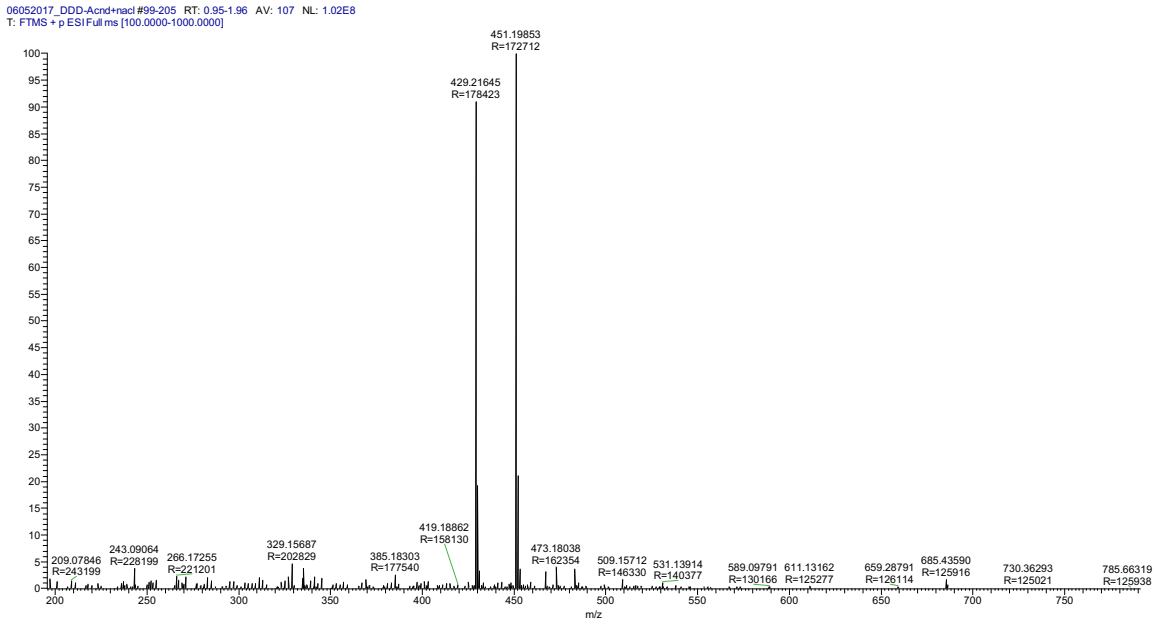
Data for DDD-12vcl. ^{1}H NMR (400 MHz, MeOD-d$_4$) δ 4.95 (m, 1H), 4.05 (m, 1H), 4.00 (m, 1H), 3.96-3.64 (m, 3H), 3.62-3.43 (m, 3H), 3.43-3.32 (m, 2H), 3.28-3.20 (m, 1H), 3.20-3.09 (m, 1H), 2.18 (m, 1H), 1.95-1.70 (m, 3H), 1.10-0.98 (m, 9H), 0.94 (d, *J* = 6.9 Hz, 3H)



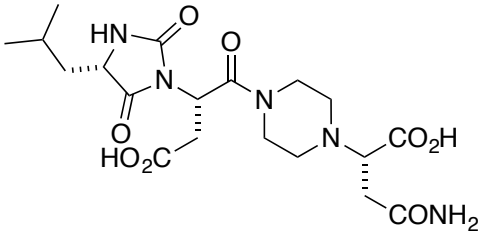
^{13}C NMR (100 MHz, MeOD-d_4) δ 175.5, 170.3, 168.2, 158.5, 67.7, 63.4, 54.9, 51.1, 43.5, 40.6, 37.1, 31.6, 26.7, 23.8, 23.7, 21.4, 19.1, 16.8



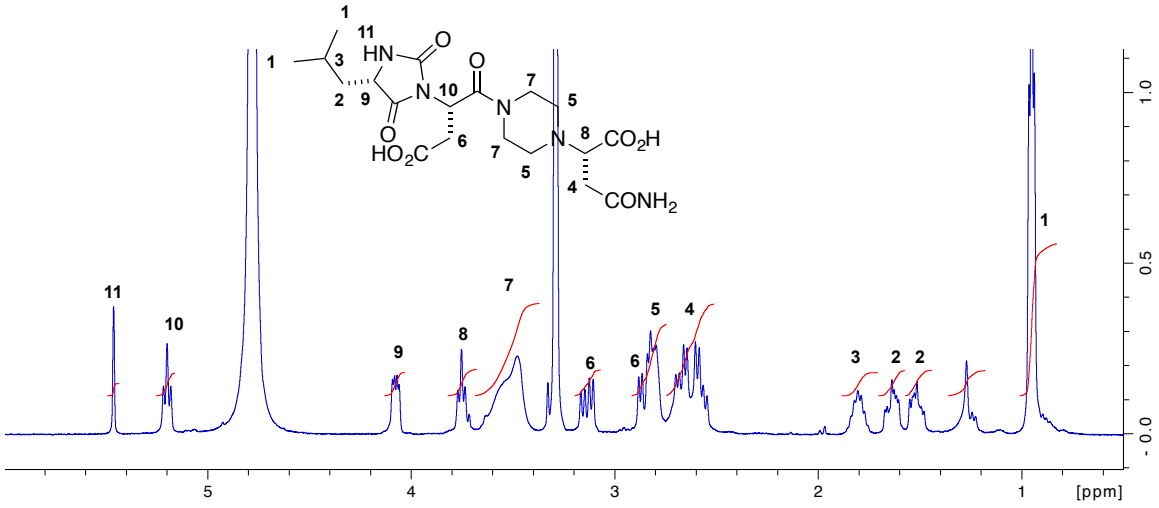
HRMS (ESI) m/z calcd for $C_{19}H_{33}N_4O_5S^+$ 429.2166; found 429.2165 $(M+H)^+$



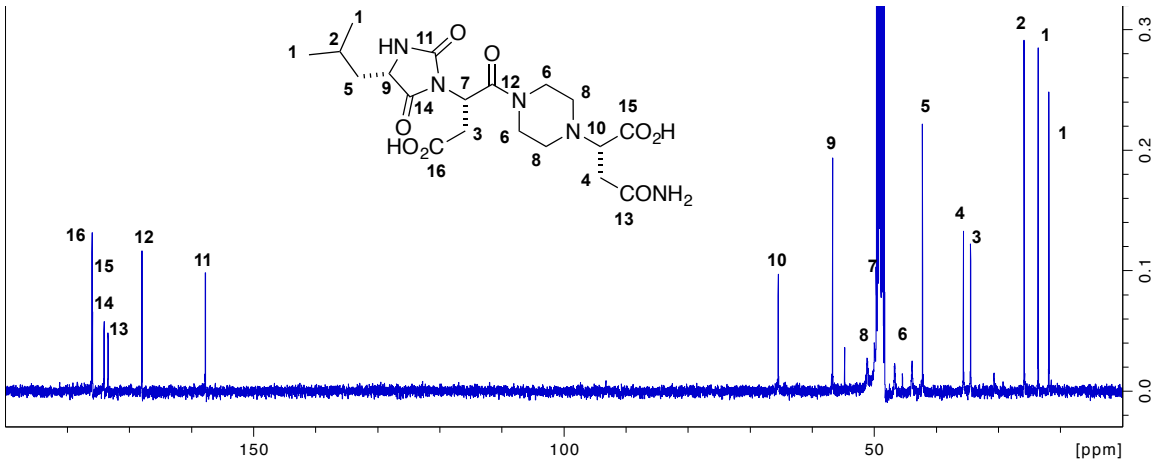
(*S*)-4-Amino-2-(4-((*S*)-3-carboxy-2-((*S*)-4-isobutyl-2,5-dioxoimidazolidin-1-yl)propanoyl)piperazin-1-yl)-4-oxobutanoic acid (LLL-12ldn)



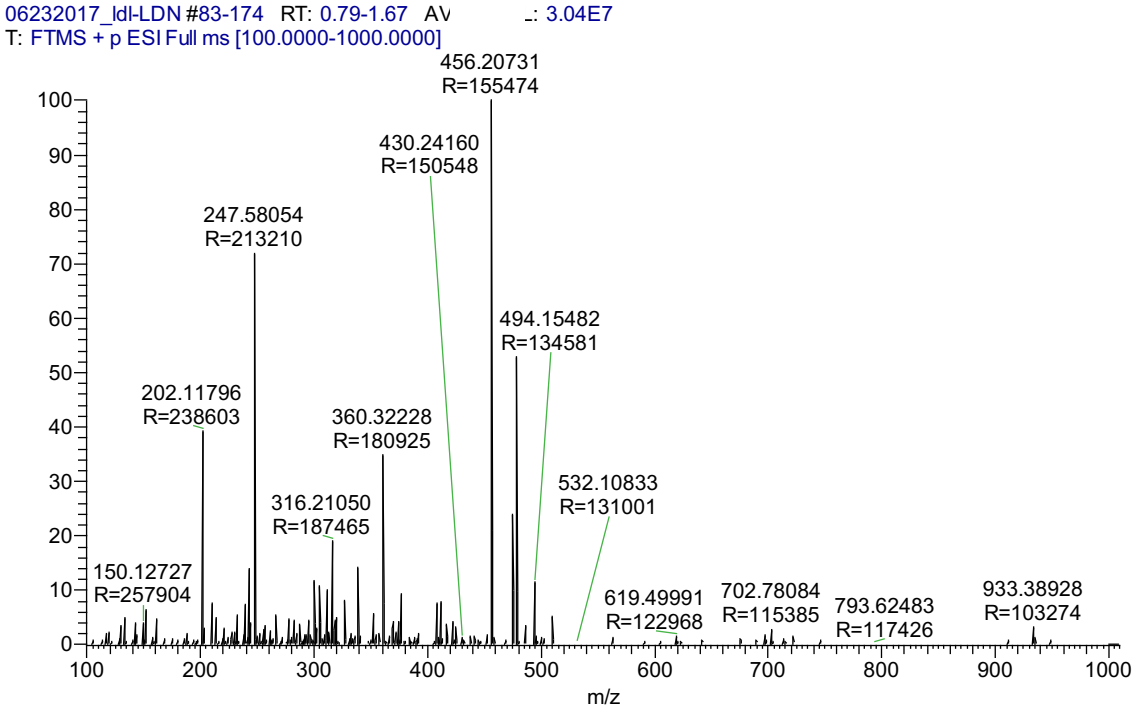
Data for LLL-12ldn. ^{1}H NMR (400 MHz, MeOD-d$_4$) δ 5.46 (s, 1H), 5.20 (t, *J* = 7.2 Hz, 1H), 4.07 (m, 1H), 3.75 (t, *J* = 6.9 Hz, 1H), 3.62-3.37 (m, 4H), 3.13 (dd, *J* = 16.4, 7.8 Hz, 1H), 2.91-2.75 (m, 3H), 2.74-2.50 (m, 4H), 1.81 (m, 1H), 1.64 (m, 1H), 1.52 (m, 1H), 0.95 (m, 6H)



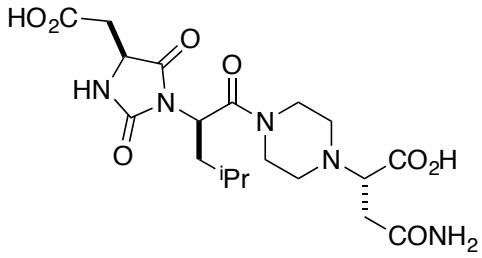
^{13}C NMR (100 MHz, MeOD-d_4) δ 176.0, 175.9, 174.1, 173.5, 168.0, 157.8, 65.5, 56.7, 51.2, 50.0, 49.7, 46.7, 43.9, 42.2, 35.6, 34.5, 25.9, 23.6, 21.9



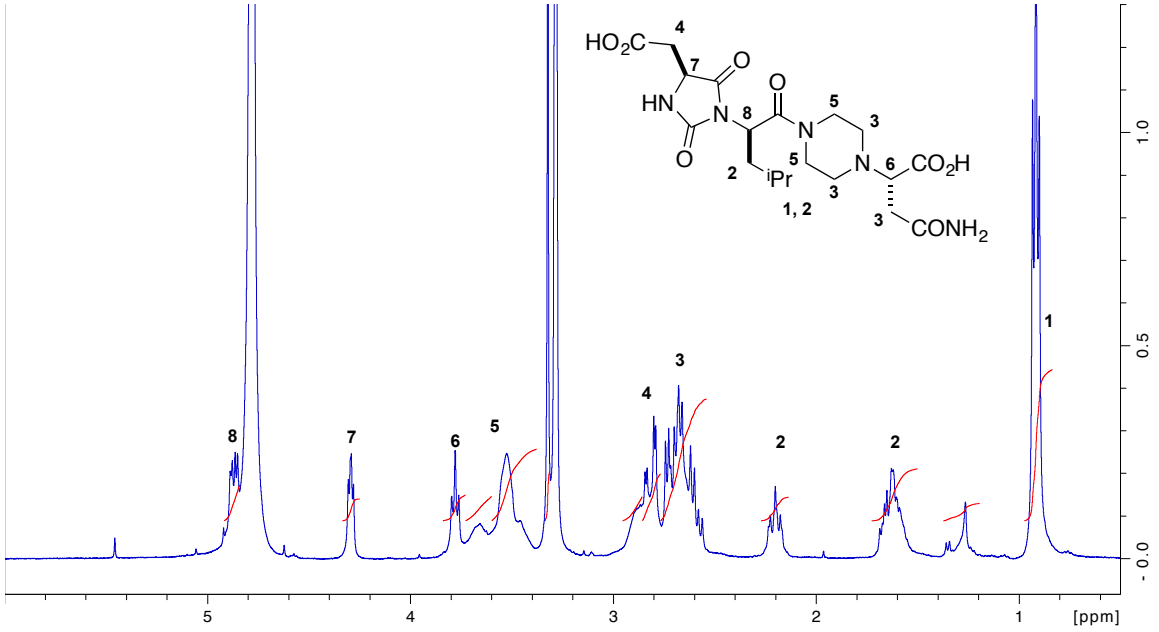
HRMS (ESI) m/z calcd for $C_{19}H_{30}N_5O_8^+$ 456.2089; found 456.2073 $(M+H)^+$



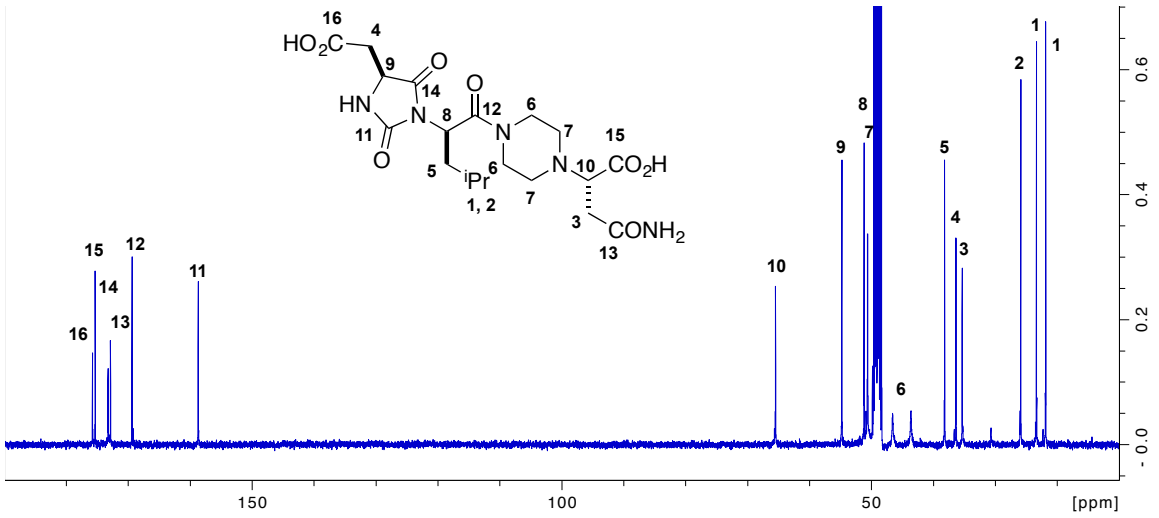
(*S*)-4-Amino-2-(4-((*R*)-2-((*S*)-4-(carboxymethyl)-2,5-dioxoimidazolidin-1-yl)-4-methylpentanoyl)piperazin-1-yl)-4-oxobutanoic acid (LDL-12dln)



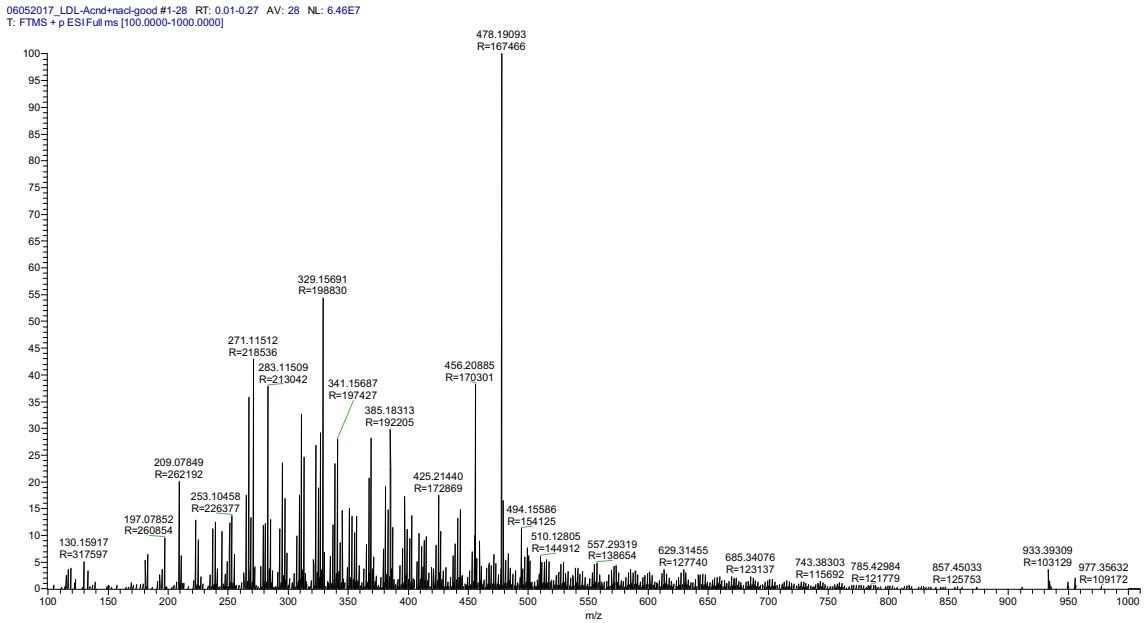
Data for LDL-12dln. ^{1}H NMR (400 MHz, MeOD-d_4) δ 4.87 (m, 1H), 4.30 (m, 1H), 3.78 (m, 1H), 3.72-3.37 (m, 4H), 2.95-2.76 (m, 3H), 2.77-2.54 (m, 5H), 2.20 (m, 1H), 1.63 (m, 2H), 0.92 (m, 6H)



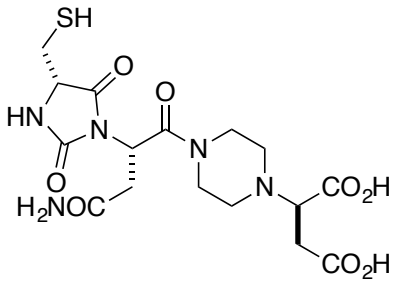
^{13}C NMR (100 MHz, MeOD-d_4) δ 175.8, 175.3, 173.3, 172.9, 169.4, 158.7, 65.5, 54.8, 51.3, 50.7, 46.6, 43.6, 38.2, 36.4, 35.3, 25.9, 23.4, 21.9



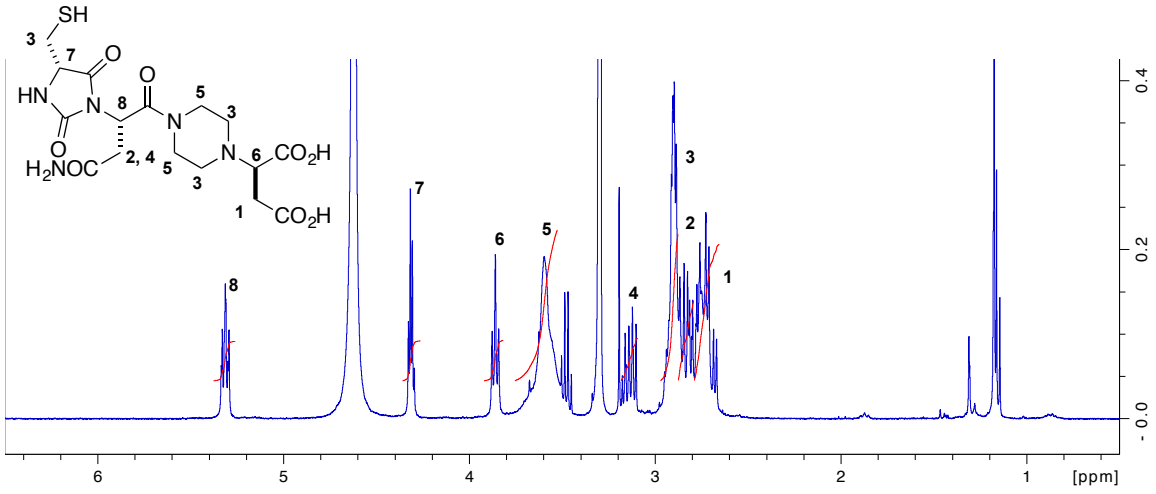
HRMS (ESI) m/z calcd for $C_{19}H_{30}N_5O_8^+$ 456.2089; found 456.2089 $(M+H)^+$



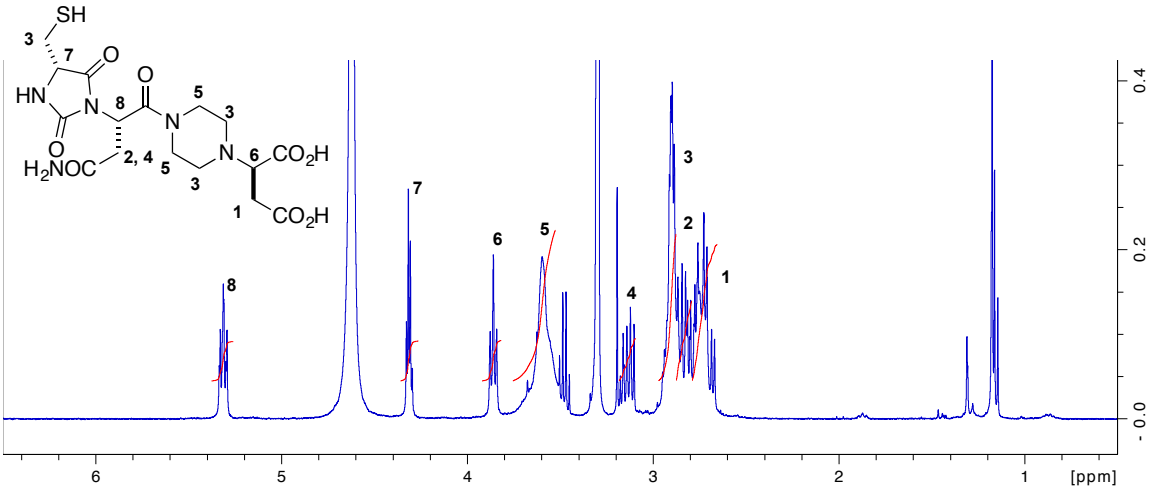
(*R*)-2-(4-((*S*)-4-Amino-2-((*S*)-4-(mercaptomethyl)-2,5-dioxoimidazolidin-1-yl)-4-oxobutanoyl)piperazin-1-yl)succinic acid (DLD-12cnd)



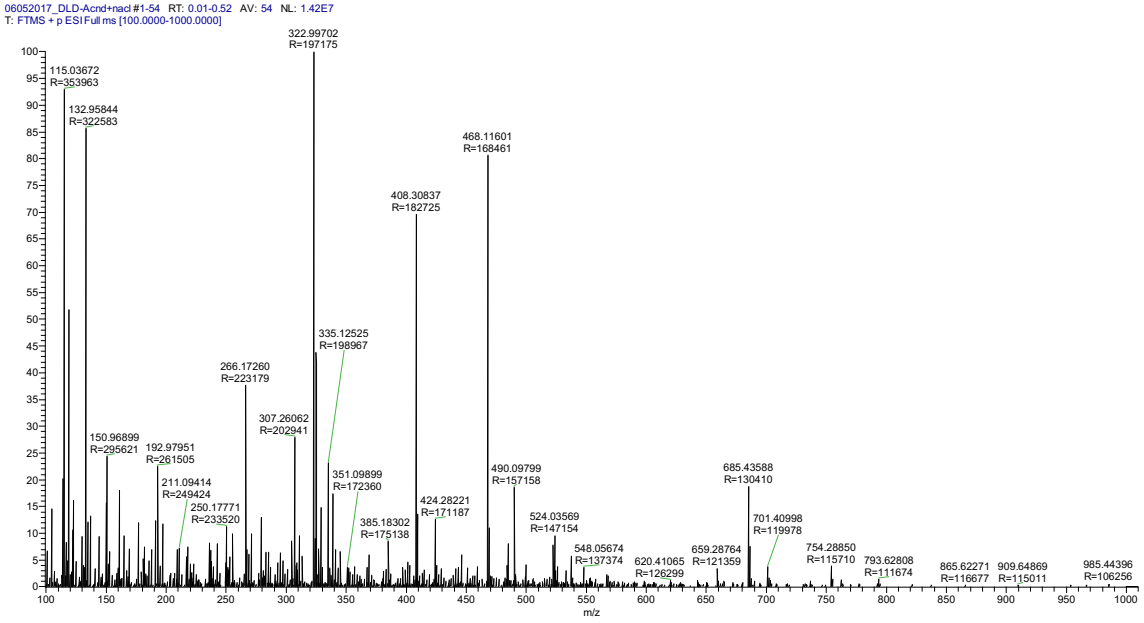
Data for DLD-12cnd. ^{1}H NMR (400 MHz, MeOD-d_4) δ 5.32 (m, 1H), 4.31 (m, 1H), 3.86 (t, *J* = 7.0 Hz, 1H), 3.75-3.52 (m, 4H), 3.14 (m, 1H), 2.97-2.87 (m, 4H), 2.87-2.79 (m, 2H), 2.79-2.65 (m, 3H)



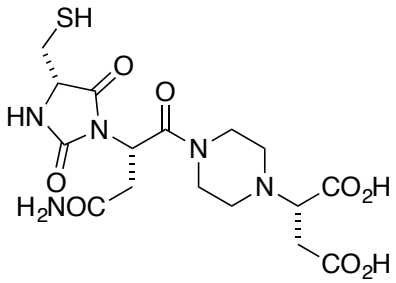
^{13}C NMR (100 MHz, MeOD-d_4) δ 174.6, 174.3, 174.0, 172.6, 168.0, 158.1, 65.1, 59.4, 50.8, 49.1, 46.4, 43.6, 35.4, 34.4, 26.1



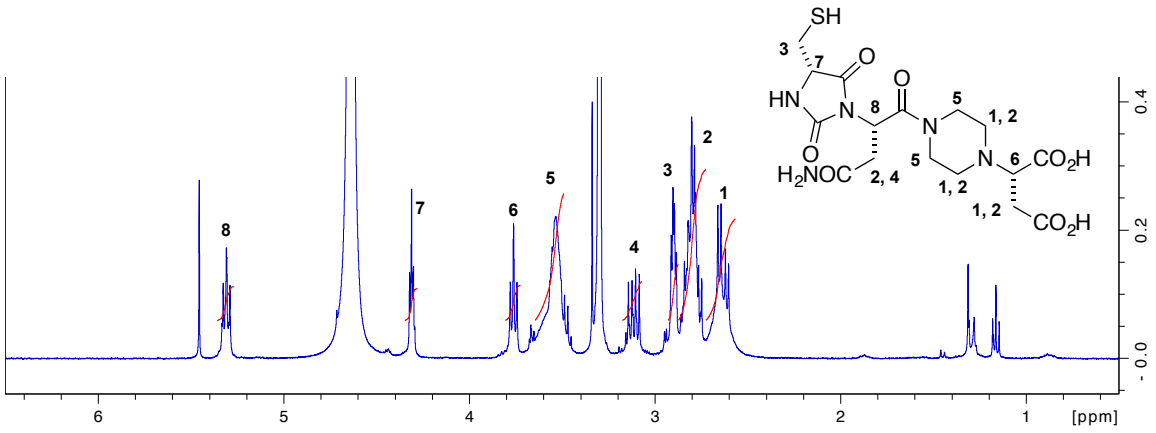
HRMS (ESI) m/z calcd for $C_{16}H_{23}N_5O_8SNa^+$ 468.1160; found 468.1160 $(M+Na)^+$



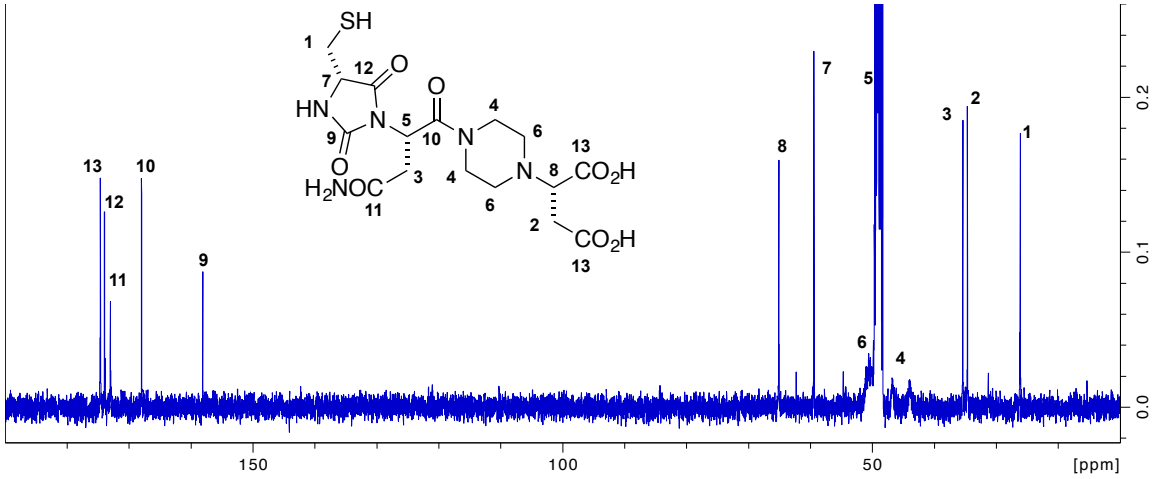
(*S*)-2-(4-((*S*)-4-Amino-2-((*S*)-4-(mercaptomethyl)-2,5-dioxoimidazolidin-1-yl)-4-oxobutanoyl)piperazin-1-yl)succinic acid (DLL-12cnd)



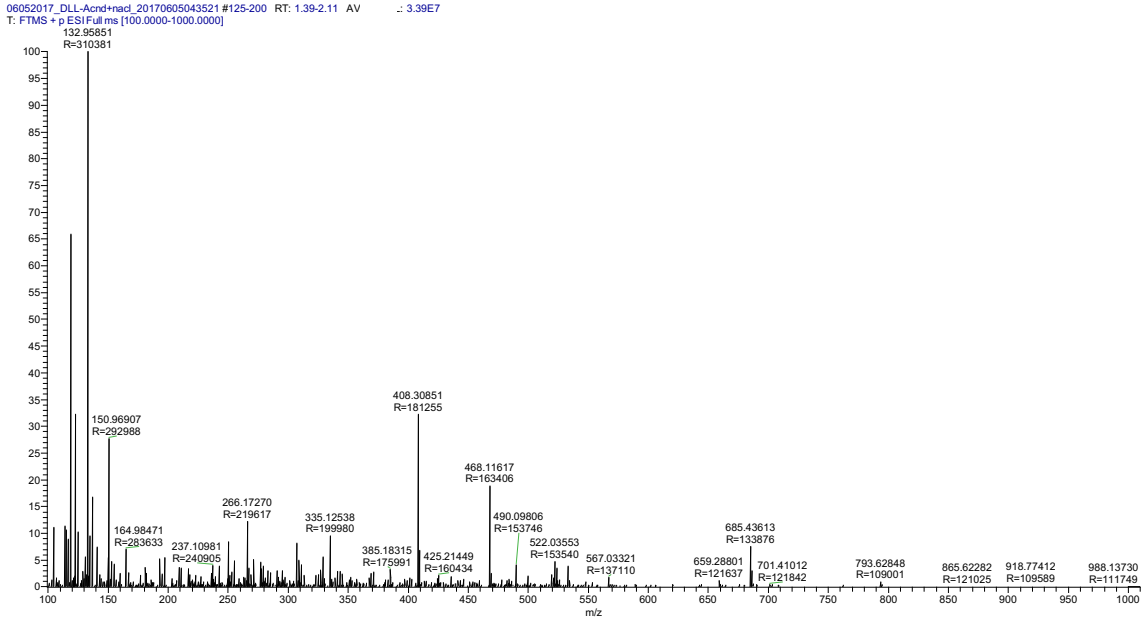
Data for DLL-12cnd. ^{1}H NMR (400 MHz, MeOD-d_4) δ 5.31 (m, 1H), 4.31 (m, 1H), 3.76 (t, *J* = 7.3 Hz, 1H), 3.69-3.45 (m, 4H), 3.11 (dd, *J* = 15.5, 7.9 Hz, 1H), 2.90 (m, 2H), 2.87-2.73 (m, 4H), 2.72-2.56 (m, 3H)



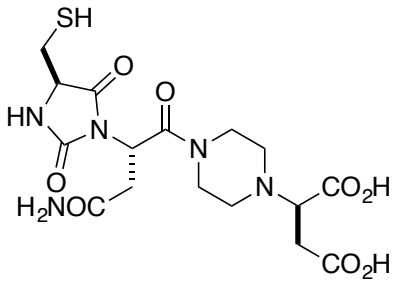
^{13}C NMR (100 MHz, MeOD-d_4) δ 174.6, 174.0, 168.0, 158.1, 65.1, 59.4, 50.6, 48.9, 46.8, 44.0, 35.4, 34.7, 26.1



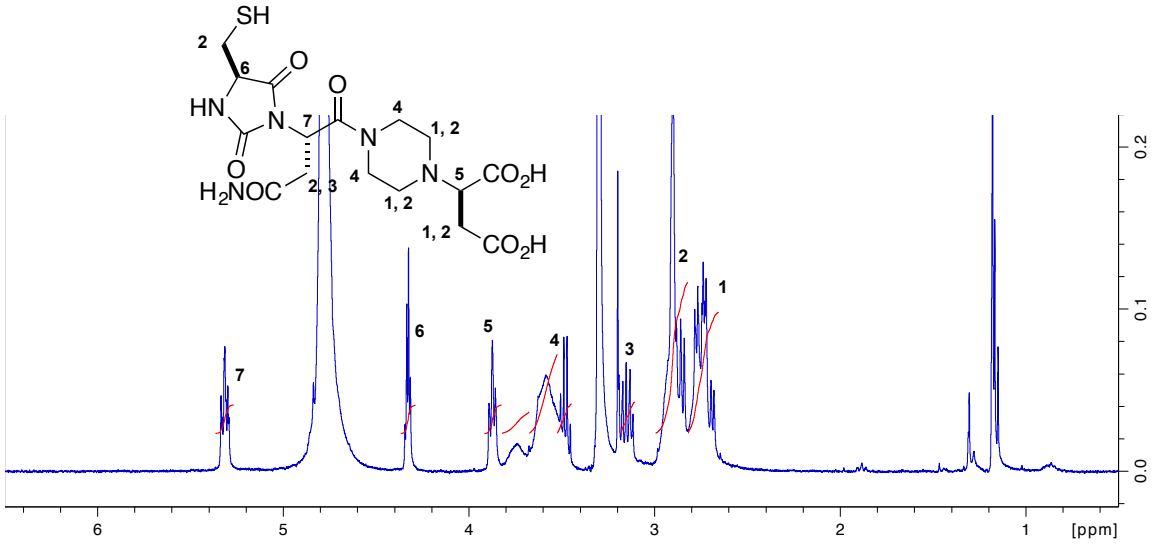
HRMS (ESI) m/z calcd for $C_{16}H_{23}N_5O_8SNa^+$ 468.1160; found 468.1162 $(M+Na)^+$



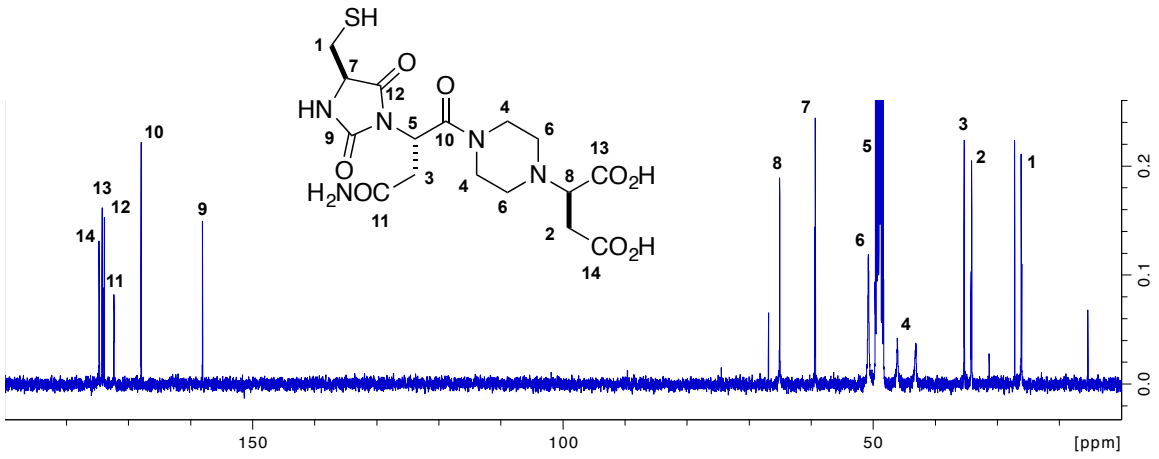
(*R*)-2-(4-((*S*)-4-Amino-2-((*R*)-4-(mercaptomethyl)-2,5-dioxoimidazolidin-1-yl)-4-oxobutanoyl)piperazin-1-yl)succinic acid (LLD-12cnd)



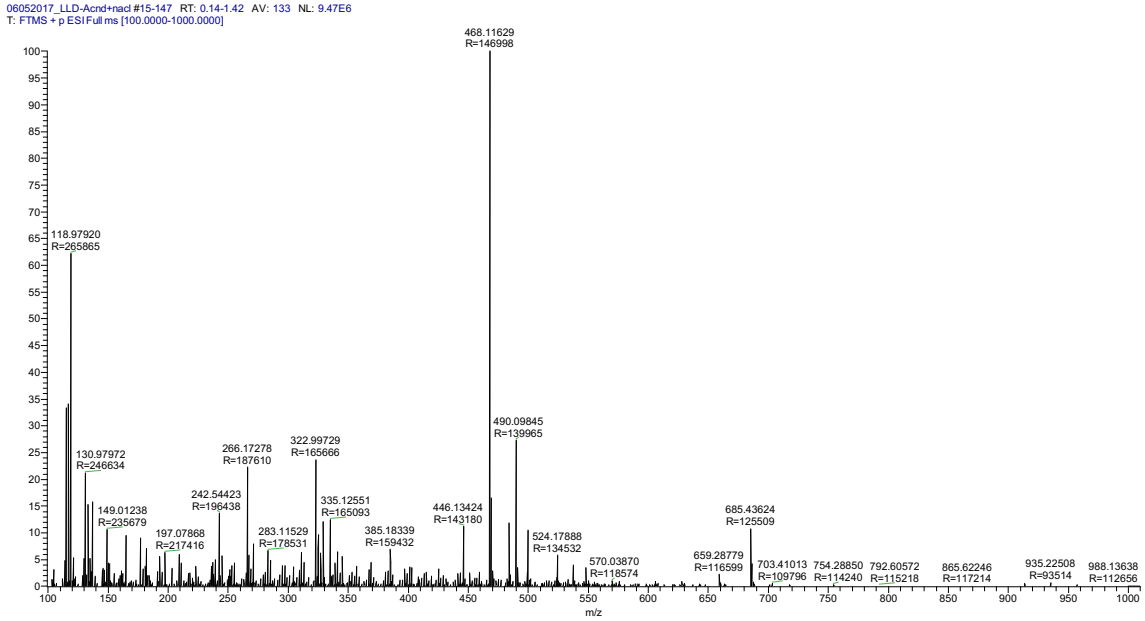
Data for LLD-12cnd. ^{1}H NMR (400 MHz, MeOD-d_4) δ 5.31 (m, 1H), 4.33 (m, 1H), 3.88 (t, *J* = 6.9 Hz, 1H), 3.82-3.53 (m, 4H), 3.15 (m, 1H), 3.00-2.82 (m, 5H), 2.82-2.65 (m, 4H)



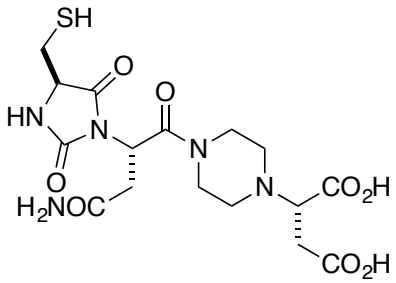
^{13}C NMR (100 MHz, MeOD-d_4) δ 174.8, 174.2, 173.9, 172.3, 168.0, 158.1, 66.9, 65.1, 59.4, 50.8, 48.8, 46.1, 43.1, 35.3, 34.1, 26.2



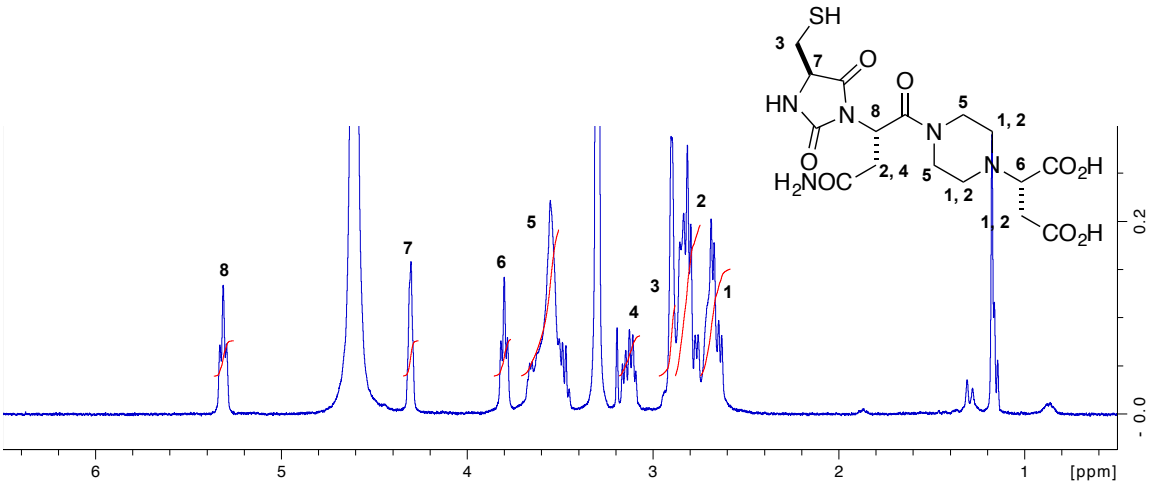
HRMS (ESI) m/z calcd for $C_{16}H_{23}N_5O_8SNa^+$ 468.1160; found 468.1163 $(M+Na)^+$



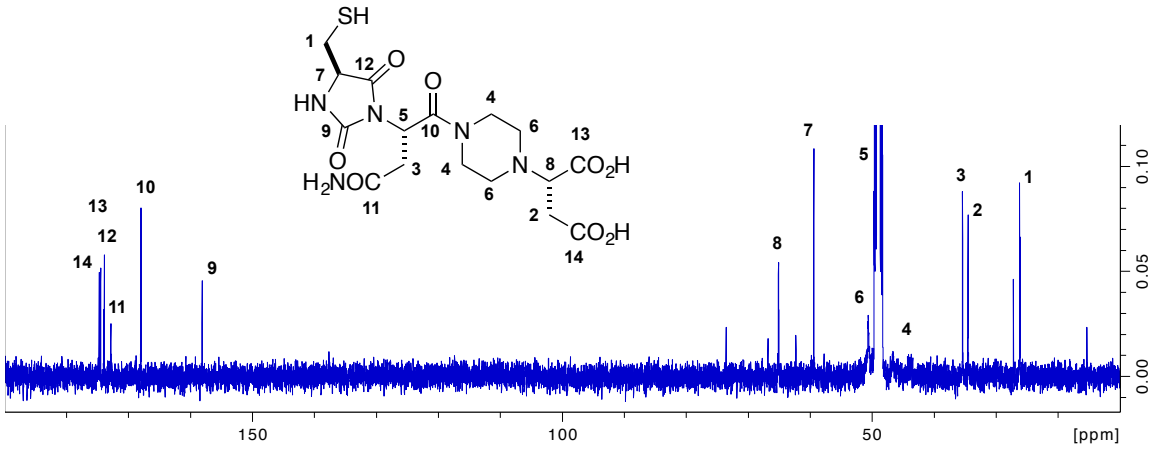
(*S*)-2-(4-((*S*)-4-Amino-2-((*R*)-4-(mercaptomethyl)-2,5-dioxoimidazolidin-1-yl)-4-oxobutanoyl)piperazin-1-yl)succinic acid (LLL-12cnd)



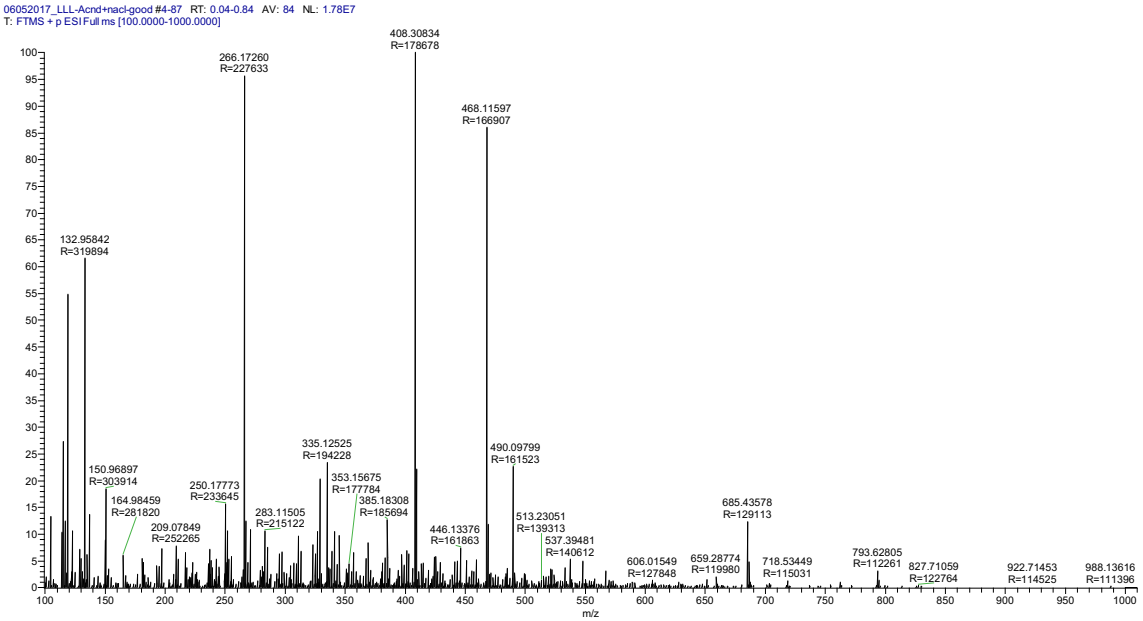
Data for LLL-12cnd. ^{1}H NMR (400 MHz, MeOD-d_4) δ 5.31 (t, *J* = 7.1 Hz, 1H), 4.31 (m, 1H), 3.80 (t, *J* = 7.0 Hz, 1H), 3.71-3.50 (m, 4H), 3.14 (m, 1H), 2.97-2.88 (m, 2H), 2.88-2.75 (m, 4H), 2.75-2.58 (m, 3H)



^{13}C NMR (100 MHz, MeOD-d_4) δ 174.7, 174.5, 173.9, 172.9, 168.0, 158.1, 65.1, 59.4, 50.6, 48.9, 46.6, 43.8, 35.4, 34.5, 27.2, 26.2



HRMS (ESI) m/z calcd for $C_{16}H_{23}N_5O_8SNa^+$ 468.1160; found 468.1160 $(M+Na)^+$



Determination of PCSK9•LDLR Inhibitory Efficiency by Time-Resolved Förster Resonance Enengy Transfer (TR-FRET) Assay

Compounds **12** were screened for their inhibitory efficiencies using PCSK9•LDLR TR-FRET assay kit (BPS Bioscience, CA) following manufacture's protocol. LDLR-Eu (125 ng/mL), dye-labeled acceptor and assay buffer were added into 384-well microplate. Compounds **12**, Pep2-8 and vehicle at various concentrations were added, and the reactions were initiated by adding biotinylated PCSK9 (2 μg/mL). Plates were shaken at room temperature for 2 h. TR-FRET fluorescent intensities were measured using plate reader with two excitation/emission wavelengths (320/620 and 320/665 nm) with 60 μs lag time and 500 μs integration time. Results were analyzed by normalizing TR-FRET emission ratio (665/620 nm) of vehicle with biotinylated PCSK9 as 0% (negative; no inhibition) and vehicle without biotinylated PCSK9 as 100% (positive; no interaction). FRET intensities of each compound were scaled according to those controls.

Table C.1 *QikProp calculations of compounds **12***

Compounds	QPlogP$_{o/w}$	PSA (Å^3)	QPPCaco (nm/s)	Rule of Five	Rule of Three
DLL-**12ncl**	-2.30	188	0.51	1	2
DDD-**12vcl**	-0.04	137	10.80	0	1
LLL-**12ldn**	-3.93	238	0.017	1	2
LDL-**12dln**	-3.63	241	0.01	1	2
LLL-**12cnd**	-3.92	235	0.037	1	2
DLL-**12cnd**	-3.92	235	0.037	1	2
LLD-**12cnd**	-3.92	235	0.037	1	2
DLD-**12cnd**	-3.92	235	0.037	1	2

$QlogP_{o/w}$ – predicted octanol/water partition coefficient; PSA – Van der Waals surface area of polar nitrogen and oxygen atoms; QPPCaco – predicted Caco-2 cell permeability; Rule of Five – number of violations of Lipinski's rule of five; Rule of Three – number of violations of Jorgensen's rule of three.

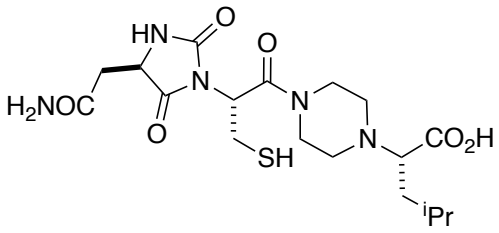
APPENDIX D

EXPERIMENTAL PROCEDURES FOR CHAPTER IV

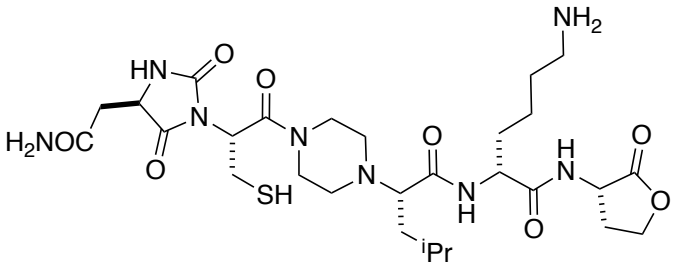
Glide

Molecular dockings of the virtual libraries were performed using Glide in the Schrödinger package (version 2015-4) [172-175] [171-174]. Selected hits from EKO analyses (Table 3.1) were used as the templates for optimizations. Side-chain substitutions and *C*-terminus modifications were enumerated using CombiGlide to obtain virtual combinatorial libraries, and those virtual structures were docked using Glide. Specifically, PCSK9 protein was obtained by deleting the LDLR protein from the same PDB file (3gcx). Alanine side-chain residues from EKO matching hits were substituted with side-chain residues corresponding to the side-chain residues from LDLR protein. Docking of virtual compounds was performed using OPLS 2005 force field within 20 Å of the grid box, and the conformations of virtual molecules were restricted within 5 Å from the parental conformers. Structures from these results represented compounds **12** docked on PCSK9 protein, and docking scores here were set as baselines. Next, another round of CombiGlide was done to substituted *C*-terminus with amino-acid lactones, in which R^4 residues were the library of *D*-, *L*- amino-acid side-chains. Docking scores from this step (compounds **16**) were compared with the previous step (compounds **12**). Top results were chosen, mostly positive functional groups, for further syntheses. However, due to the synthetic difficulty and poor stability of the hydantoin when R^1 was cysteine, this position was enumerated with other amino-acid side-chains to obtain other residues that had best docking scores, which were glutamine and glutamic derivatives.

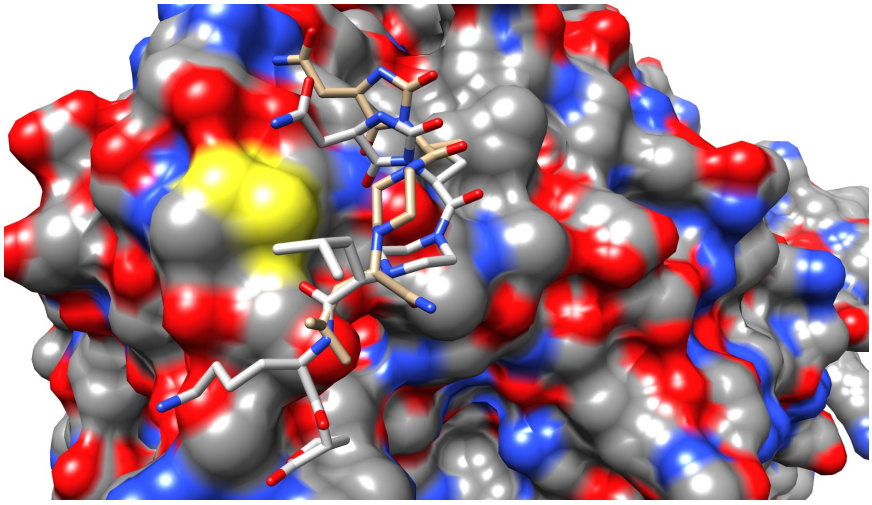
DLLD-**16nclk**



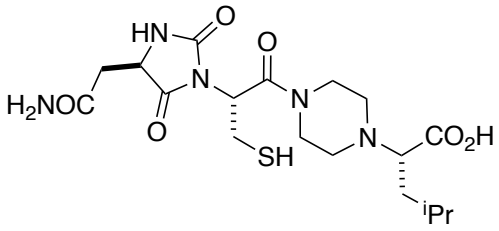
DLL-**12ncl** (gold), ΔG = -2.04 kcal/mol



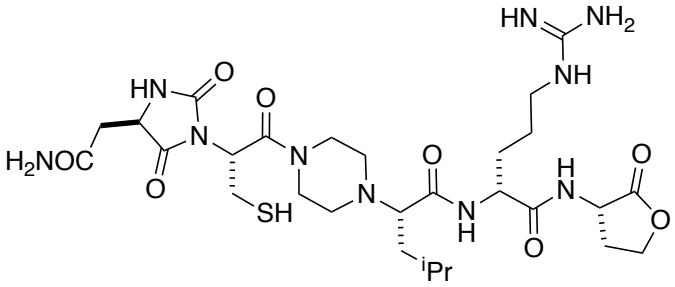
DLLD-**16nclk** (silver), ΔG = -2.31 kcal/mol



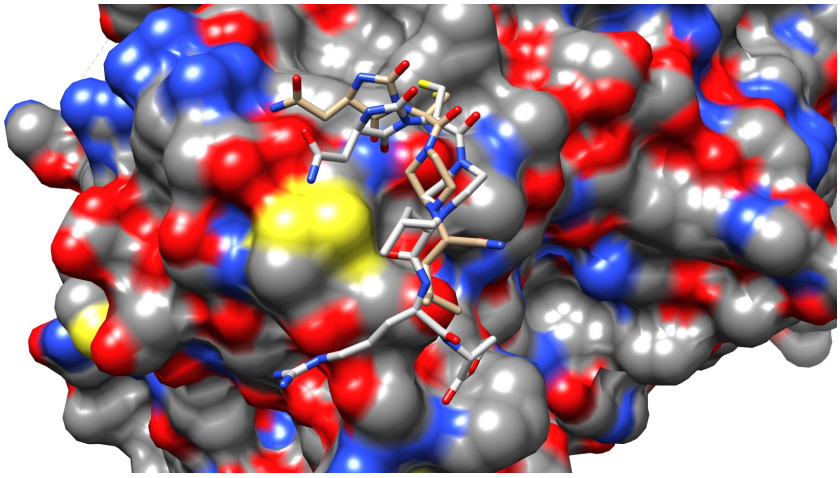
DLLD-**16nclr**



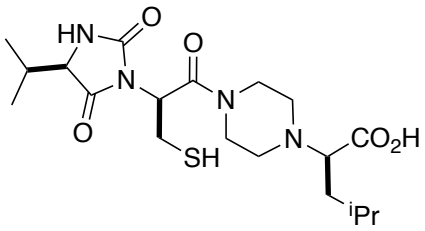
DLL-**12ncl** (gold), ΔG = -2.04 kcal/mol



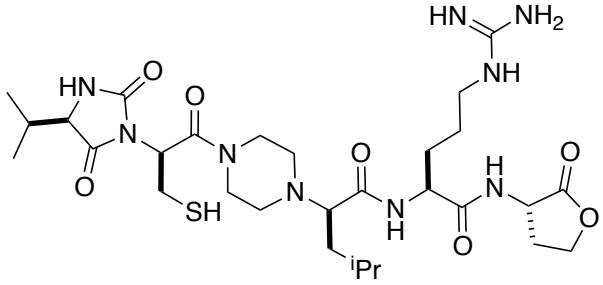
DLLD-**16nclr** (silver), ΔG = -3.41 kcal/mol



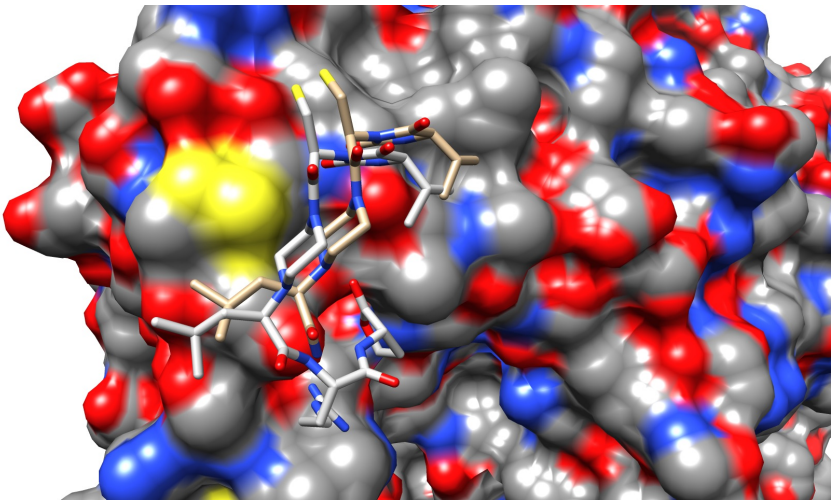
DDDL-**16vclr**



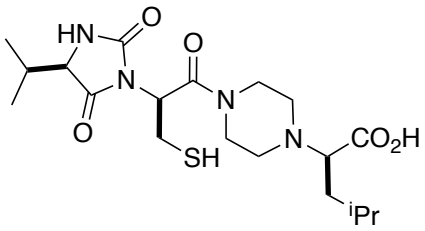
DDD-**12vcl** (gold), ΔG = -1.25 kcal/mol



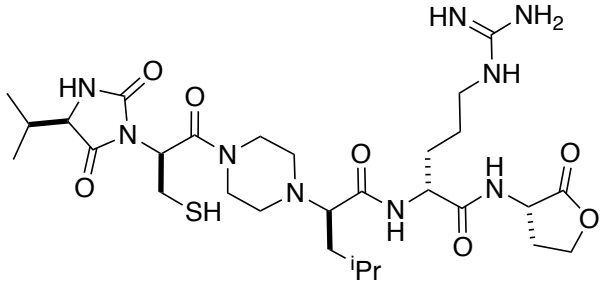
DDDL-**16vclr** (silver), ΔG = -4.88 kcal/mol



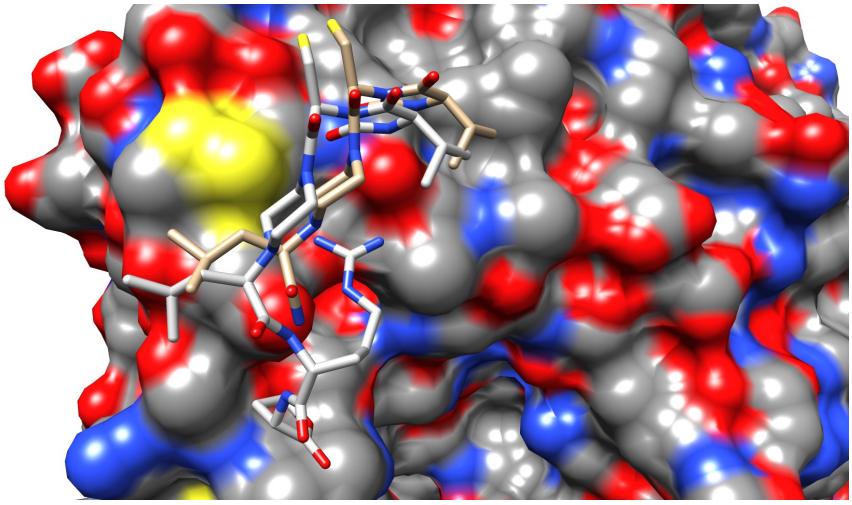
DDDD-**16vclr**



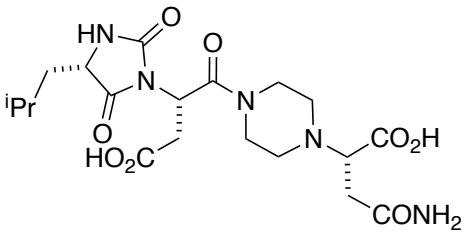
DDD-**12vcl** (gold), ΔG = -1.25 kcal/mol



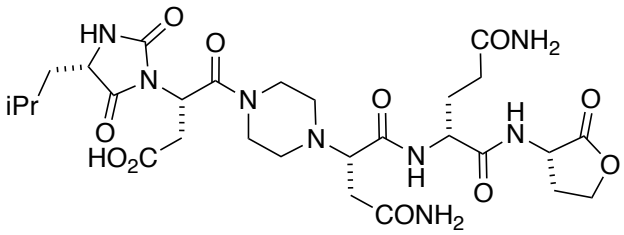
DDDD-**16vclr** (silver), ΔG = -4.46 kcal/mol



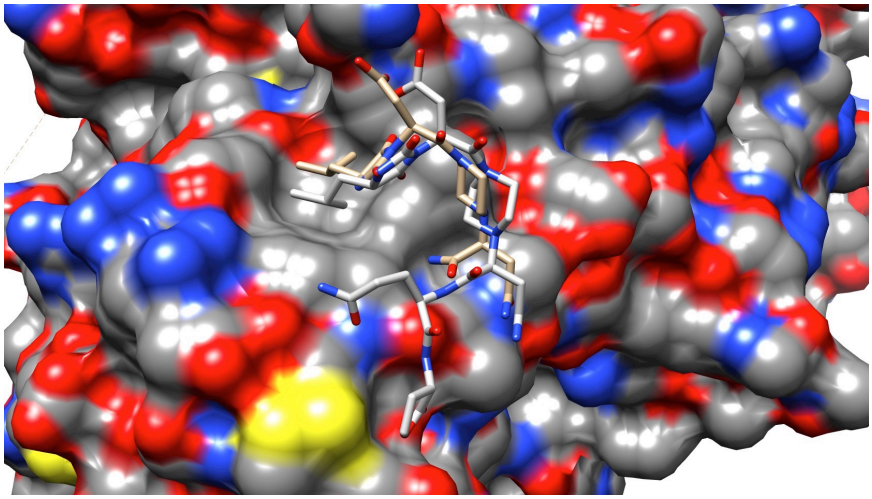
LLLD-**16ldnq**



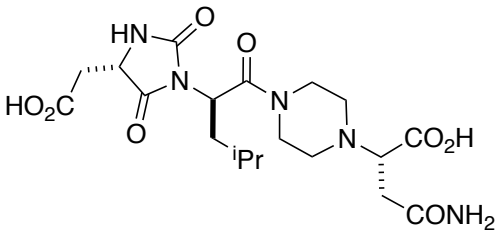
LLL-**12ldn** (gold), ΔG = -2.29 kcal/mol



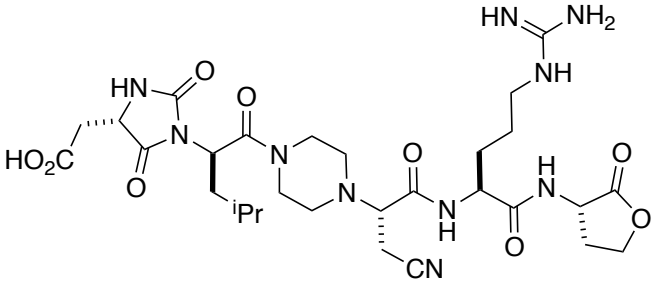
LLLD-**16ldnq** (silver), ΔG = -5.44 kcal/mol



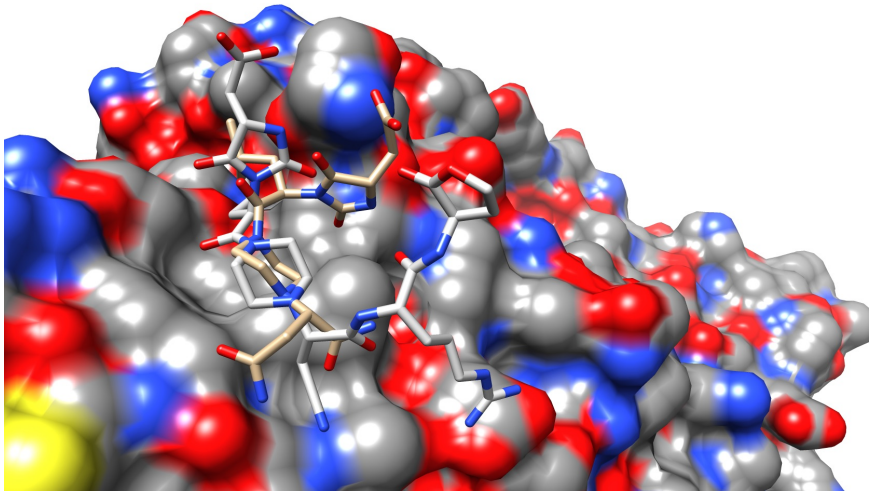
LDLL-**16dl(CN)r**



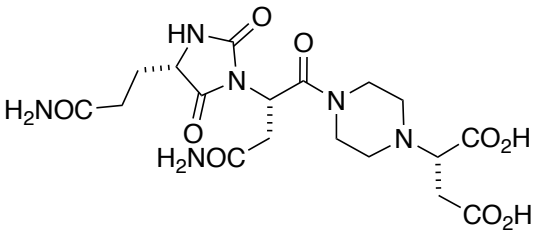
LDL-**12dln** (gold), ΔG = -1.25 kcal/mol



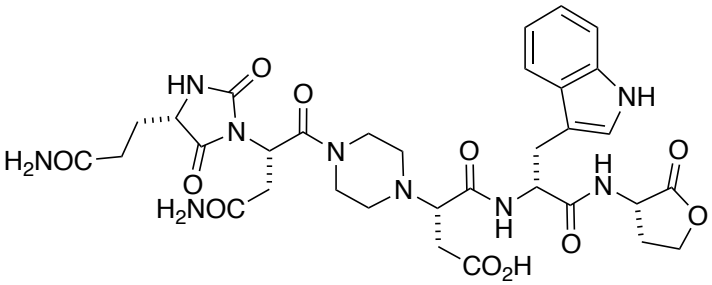
LDLL-**16dl(CN)r** (silver), ΔG = -2.67 kcal/mol



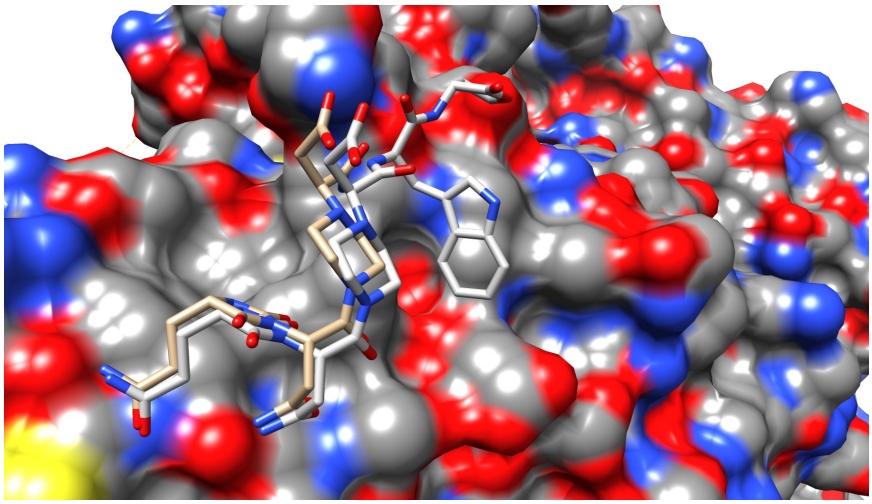
LLLD-**16qndw**



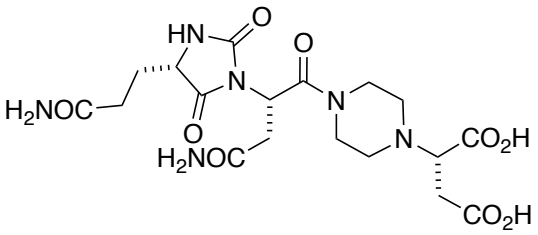
LLL-**12qnd** (gold), ΔG = -3.17 kcal/mol



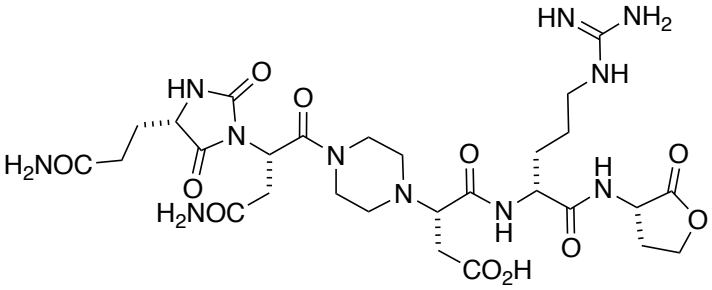
LLLD-**16qndw** (silver), ΔG = -3.63 kcal/mol



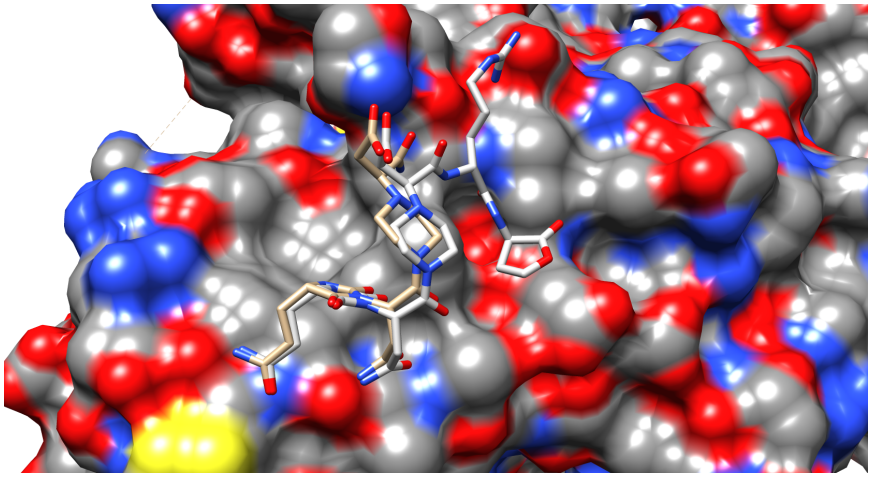
LLLD-**16qndr**



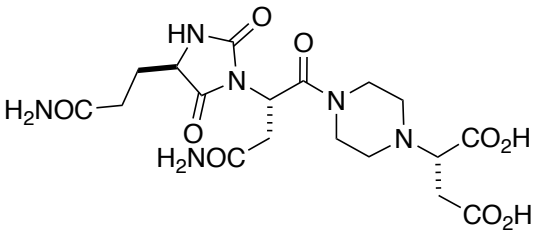
LLL-**12qnd** (gold), ΔG = -3.71 kcal/mol



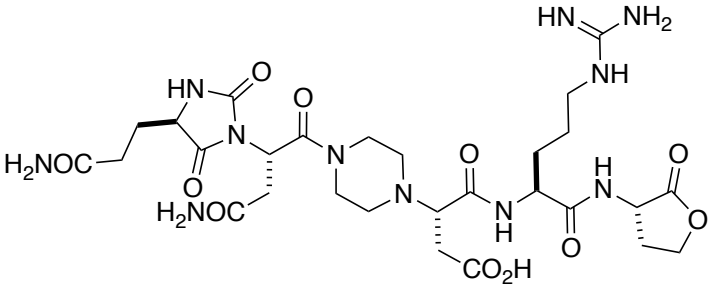
LLLD-**16qndr** (silver), ΔG = -6.15 kcal/mol



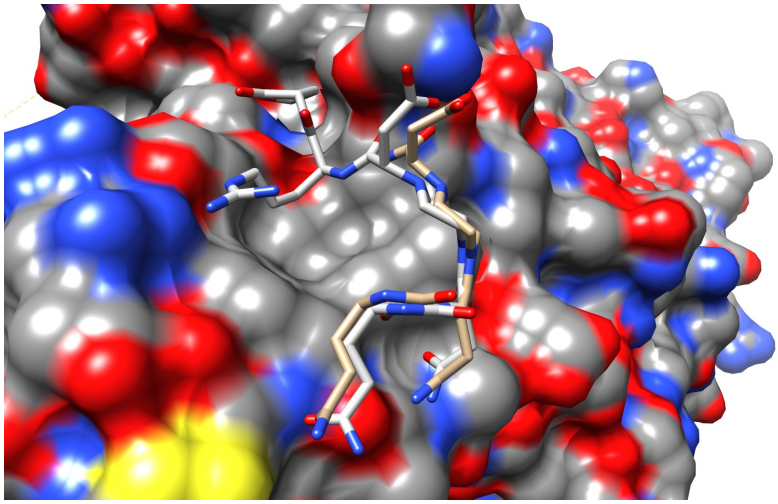
DLLL-**16qndr**



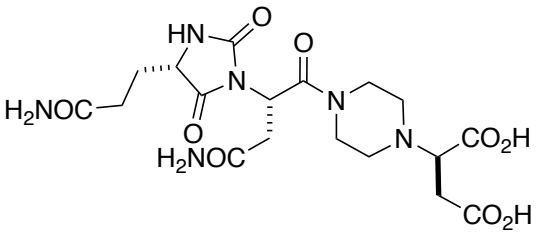
DLL-**12qnd** (gold), ΔG = -3.15 kcal/mol



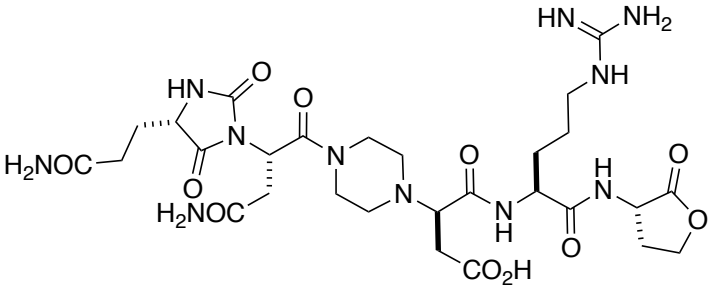
DLLL-**16qndr** (silver), ΔG = -4.77 kcal/mol



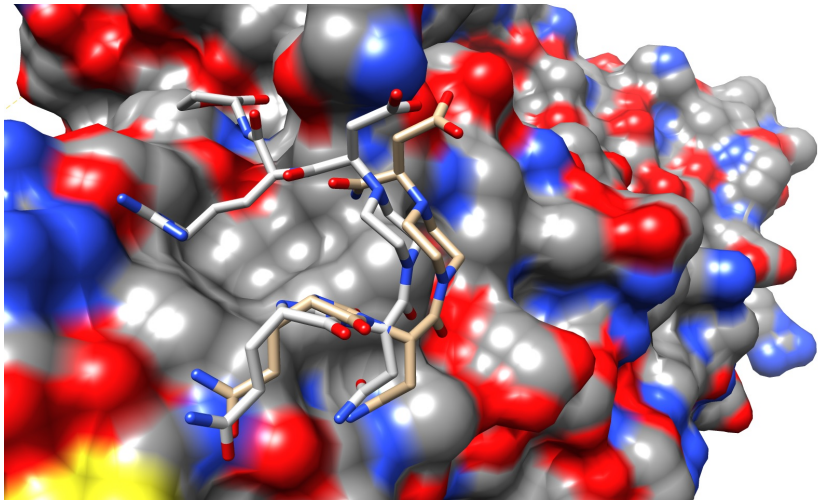
LLDL-**16qndr**



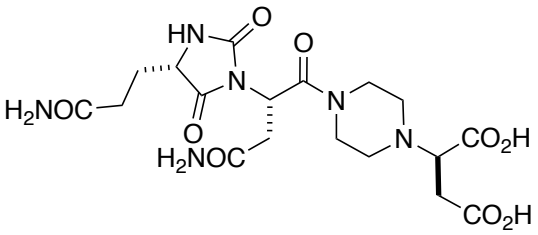
LLD-**12qnd** (gold), ΔG = -2.39 kcal/mol



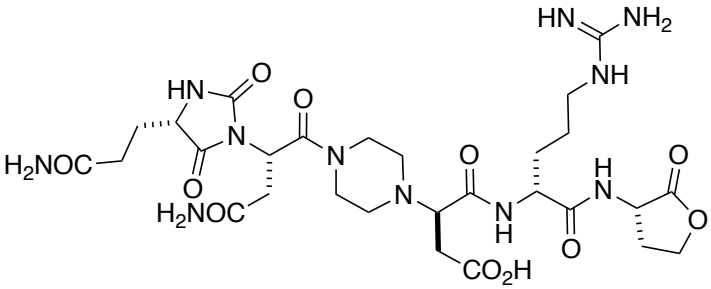
LLDL-**16qndr** (silver), ΔG = -5.21 kcal/mol



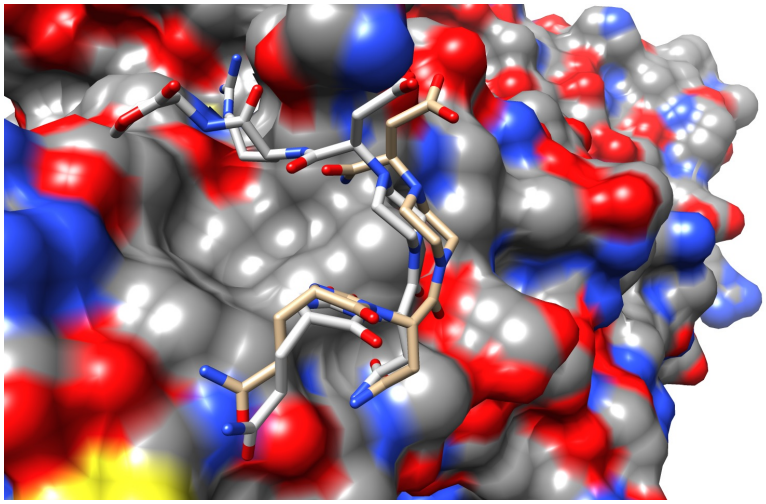
LLDD-**16qndr**



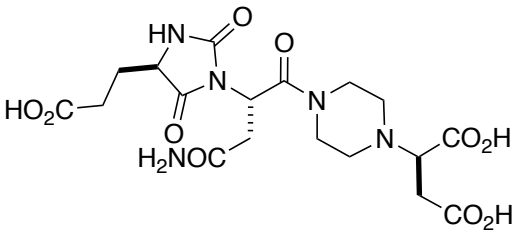
LLD-**12qnd** (gold), ΔG = -2.39 kcal/mol



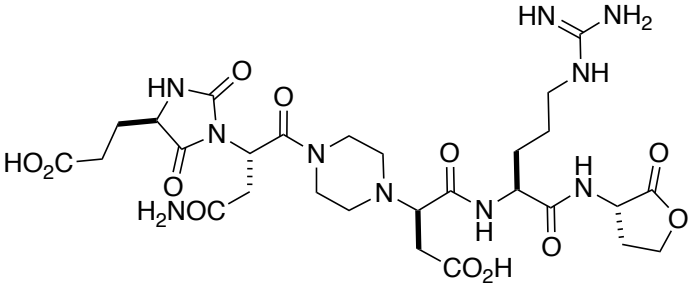
LLDD-**16qndr** (silver), ΔG = -5.37 kcal/mol



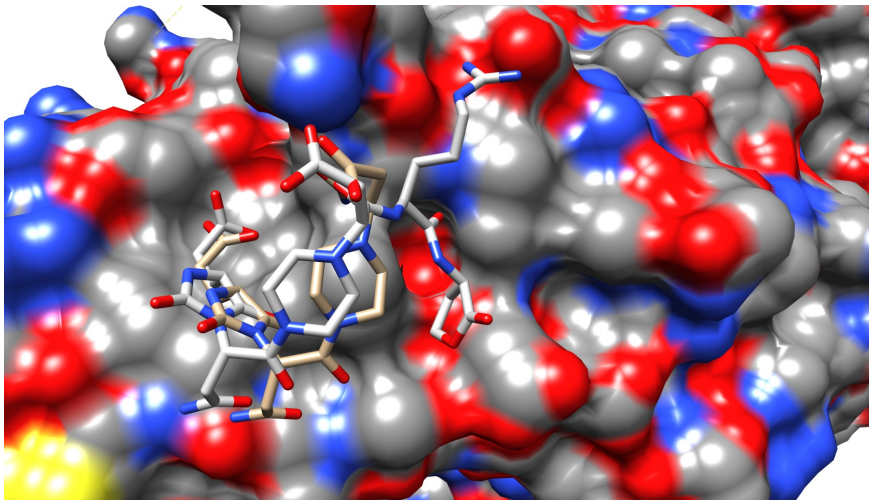
DLDL-**16endr**



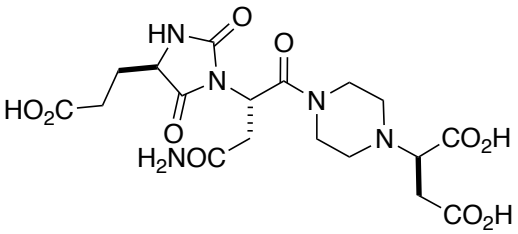
DLD-**12end** (gold), ΔG = -2.73 kcal/mol



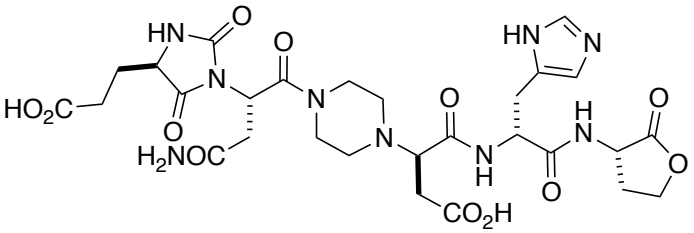
DLDL-**16endr** (silver), ΔG = -5.41 kcal/mol



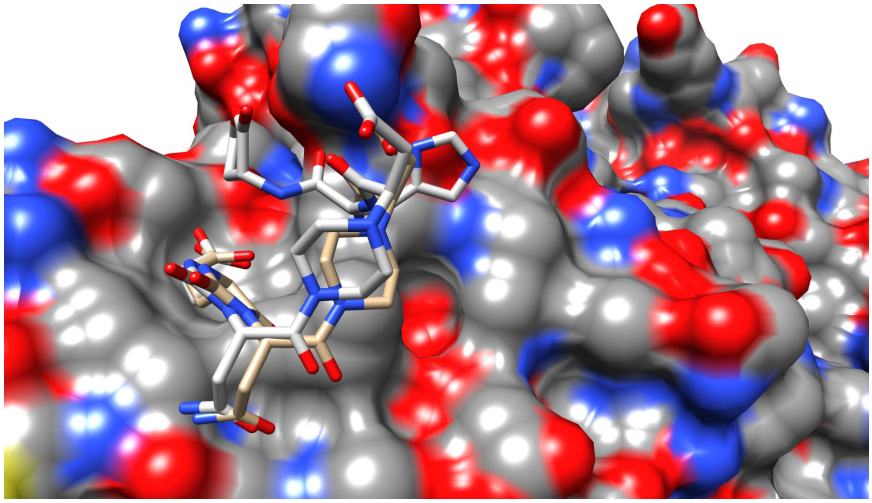
DLDD-**16endh**



DLD-**12end** (gold), ΔG = -2.73 kcal/mol



DLDD-**16endh** (silver), ΔG = -3.59 kcal/mol



LDL-uptake Assay

HepG2 cells were seeded into 96-well plate at a density of 3 x 10^4 cells/well in DMEM + 10% FBS medium (Sigma-Aldrich). After 24 h, the medium was aspirated, and cells were washed two times with PBS. Cells were treated with 10% lipoprotein-deficient serum (Gemini) in DMEM and incubated for another 24 h. After washing cells with PBS twice, 100 µL of Pep2-8 (30 µM) or inhibitors at various concentrations (50, 5, 0.5 µM), which were preincubated with 15 µg/mL PCSK9 in 10% lipoprotein-deficient/DMEM containing 0.5% DMSO for 30 minutes before use, were added, and cells were incubated for 2 h. BODIPY-LDL (10 µg/mL, Invitrogen) was added to each well, and microplate was incubated for another 3 h. Cells were washed three times with PBS, and fluorescent intensities were measured on Synergy H4 plate reader (Biotek) with excitation and emission wavelength at 488 and 520 nm, respectively.

In vitro PCSK9-LDLR Binding Assay

Direct inhibitions of compounds **16** toward PCSK9•LDLR interaction were tested using the PCSK9•LDLR *in vitro* binding assay kit (MBL Int. Co., MA) according to the manufacture's protocol. Briefly, compounds (50 µM), Pep2-8 (50 µM) or vehicle (DMSO) were preincubated with 50 ng/mL of His-tagged PCSK9 in the reaction buffer (100 µL) 30 minutes prior the experiments. The premix mixtures were added microplate containing pre-coated EGF-AB domain of LDLR protein. Microplate was incubated for 2 h at room temperature while shaking at 300 rpm on an oribital microplate shaker. Each well was washed 4 times with washing buffer (350 µL each), and the biotinylated anti-His-tag mAb was added (100 µL). Microplate was shaken at 300 rpm for 1 h at room temperature. After washing 4 times with washing buffer, HRP-conjugated Streptavidin

(100 µL) was added into each well, incubated for 20 minutes at room temperature while shaking at 300 rpm. Microplate was washed 4 times with washing buffer, then the substrate reagent (tetra-methylbenzidine, 100 µL) was added into each well, and the microplate was shaken for 20 minutes at room temperature. Finally, the reaction was stopped with 1 N H_2SO_4 (100 µL), and the absorbance was measured on plate reader at dual wavelengths of 450/540 nm.

MTT Assay

HepG2 cells were seeded at 3000 cells/well in 96-well plate, and incubated in DMEM + 10% FBS medium for 24 h. Medium containing inhibitors was added to the final concentrations of 50 µM or 100 µM (100 µL) and cells were incubated for another 2 days. MTT substrate in HBSS buffer (5 mg/mL, 20 µL) was added, and cells were incubated for additional 3 h. Medium was aspirated, then DMSO (100 µL) was added subsequently. Microplate was shaken for 5 minutes, then the absorbance was measured with plate reader at 570 nm wavelength.

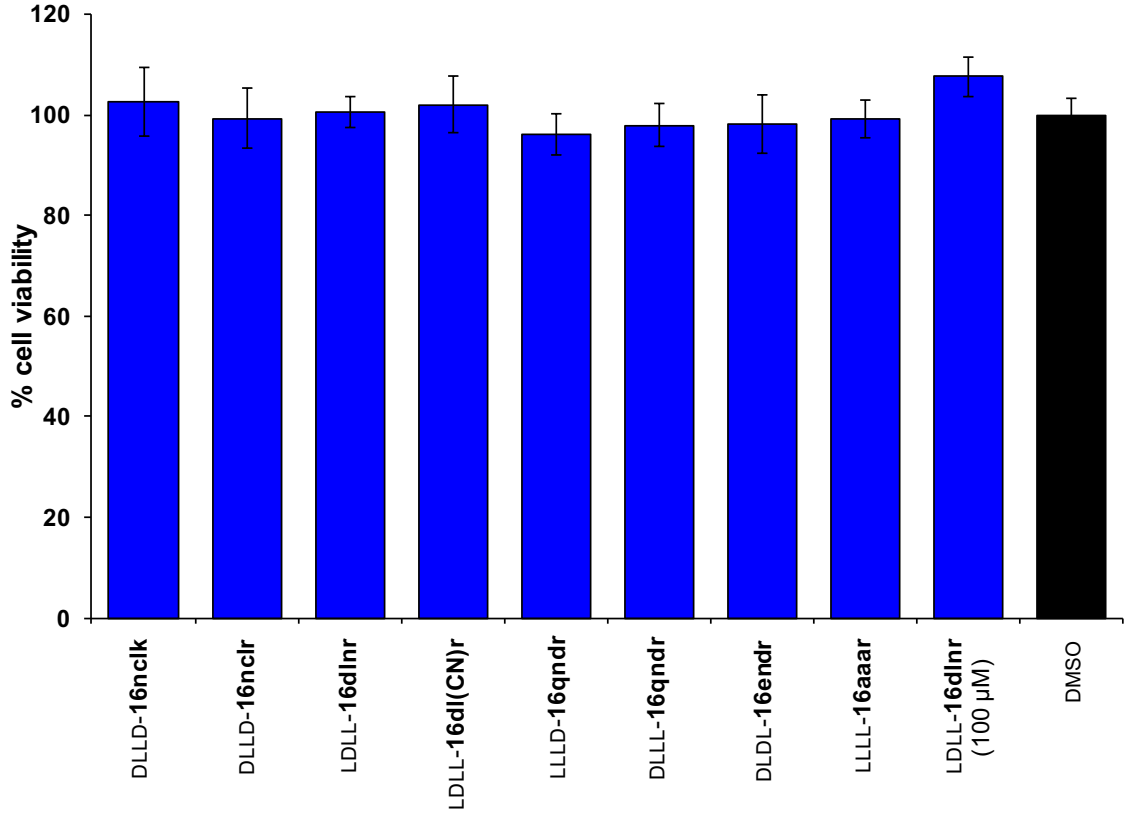


Figure D.1. Seven featured compounds **16** and one partial negative control were tested for their cytotoxicity with HepG2 cells. All compounds showed no cytotoxicity compared to the control (DMSO, black).

Determination of Water Solubilities

Compound solubilities in water were measured by UV absorbance following the literature.[196] Briefly, compounds dissolved in MeOH were plated on 96-well plate, then solvent was evaporated at room temperature. Water at various volumes were added to each well to reach at different concentrations, microplate was shaken for 5 h, and left overnight at room temperature to reach equilibrium. Microplate was centrifuged for 20 min to remove any particulates. The clear solution was transfer to a 96-well UV-transparent plate to measure the UV absorbance at 230 nm wavelength. For this study,

UV absorbances are linear beyond 100 μM for every compound, thus the solubilities of compounds are at least 100 μM.

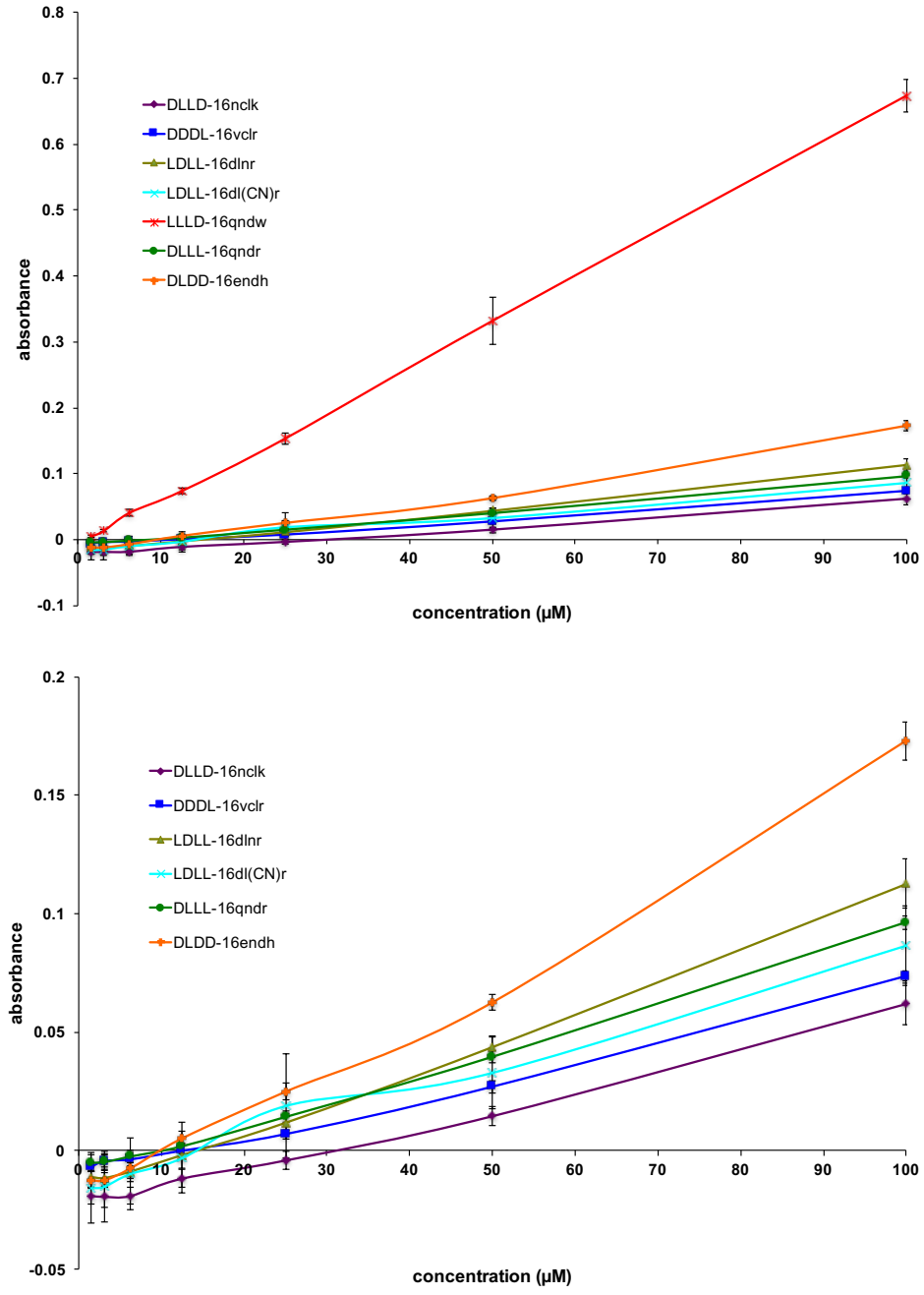


Figure D.2. Selected compounds **16** were tested for their solubility in water by UV absorbance method. Compounds at seven concentrations up to 100 μM gave linear absorbance trend, which implied the solubility of compounds were at least > 100 μM concentration.

Photoaffinity Labeling

Compounds **18** (100, 200 and 300 μM) were incubated with 0.25 μg/μL PCSK9 in 100 mM HEPES buffer to a final volume of 20 μL at 4 ºC for 1.5 h. Mixtures were transferred into 96-well plate, and then irradiated by a LED UV flashlight (365 nm) at 4 ºC for 30 min. Samples (19.5 μL) were transferred into Eppendorf tubes, mixed with 2.5 μL of 10% SDS, and added 0.5 μL of 10 mM azide-fluor 488 (Sigma-Aldrich) to each tube. Huisgen cycloaddition was initiated by adding 2.5 μL of catalyst mixture, which was prepared immediately before use by mixing 3:1:1 v/v of 1.7 mM TBTA in 80 % *tert*-butanol/20% DMSO, 50 mM $CuSO_4$ and 50 mM TCEP. Reaction was kept at room temperature for 4 h. Finally, the reactions were quenched with 5 μL of 6x SDS buffer, and boiled at 95 ºC for 10 min to denature PCSK9 protein. SDS-PAGE was performed by loading on 10% polyacrylamide gel. The in-gel fluorescence was determined by Typhoon FLA 9500 imager at 488 nm wavelength, and all proteins were stained with Coomassie Brilliant Blue G250. For the competitive experiment, 50-fold concentration of compounds **17** were incubated with PCSK9 at 4 ºC for 1 h before adding compounds **18**. Also, the mixture of compound **18** and PCSK9 without irradiating UV flashlight (negative control) had been tested to gauge non-specific binding.

Surface Plasmon Resonance (SPR)

Direct binding of each compound to PCSK9 was measured by SPR using a Biacore T200 optical biosensor (GE Healthcare) at 25˚C. CM5 sensor chips (GE Healthcare) were used to create PCSK9 biosensors using standard amine coupling chemistry. Briefly, after mixing equal volume mixture of 0.1 M N-hydroxysuccinimide and 0.4 M 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide, the mixture was injected for 7 min at 10

µl/min to activate the flow cell surface. PCSK9 stock (about 1 mg/ml) was diluted to 10 µg/ml in 50 mM sodium acetate (pH 5.5) and injected to the activated surface and immediately followed by a 5-min injection of 1 M ethanolamine (pH 9) to deactivate the surface. By varying the ligand contact times for 5, 7 or 10 minutes, different density of PCSK9 surfaces were generated: 2600, 4300 and 5400 RU respectively. Phosphate buffered saline (PBS: 8.06 mM Na_2HPO_4 and 1.94 mM KH_2PO_4, pH 7.4, 2.7 mM KCl, 137 mM NaCl) was used as running buffer for immobilization. A reference flow cell was prepared with activation and deactivation steps but no protein coupled. All binding experiments were performed in a running buffer of 50 mM Tris-HCl (pH 7.5), 150 mM NaCl, and 5% (v/v) DMSO (TBSD) using a flow rate of 50 µl/min.

To evaluate compound binding, each compound was dissolved in 100% DMSO to 10 mM and then diluted 20-fold in TBS to 500 µM final concentration as working stock for further dilution in TBSD. Samples were injected over the PCSK9 surface for 30 s followed by 60 s of dissociation and a subsequence wash step with 50% DMSO solution. In the screening experiment, the positive control Pep2-8 was injected three times (beginning, middle and end) during the run. The signal 10 s before injection stop of the sensorgram was treated as PCSK9-binding response. The SPR sensorgrams were solvent corrected, reference and buffer subtracted, and evaluated using the Biacore T200 Evaluation Software (version 3.1).

Table D.1. Confirmation of compound binding and their dissociation constants for PCSK9 as determined by SPR.

Compound	k_a (x 10^3 $M^{-1}s^{-1}$)	k_d (x 10^{-2} s^{-1})	K_D^{Kin} (μM)	%R_{max}	K_D^{SS} (μM)
DLLD-**16nclk**	2.99 (0.89)	10.2 (1.1)	41.2 (17.5)	30 (13)	303 (94)
LDLL-**16dlnr**	4.04 (2.20)	8.74 (3.40)	24.8 (9.1)	70 (24)	55.4 (13.2)
LDLL-**16dl(CN)r**	2.50 (1.44)	8.06 (3.07)	35.8 (11.4)	59 (12)	74.5 (15.6)
Pep2-8	435 (202)	156 (78)	3.56 (0.16)	32 (11)	7.63 (1.42)

Data are average of three independent experiments using different PCSK9 surface density (R_{PCSK9} = 2600 RU, 4300 RU, or 5400 RU) with standard deviations shown in parentheses. The theoretical maximal binding response of each surface was calculated using equation: R_{max} = ($MW_{compound}/MW_{PCSK9}$) x R_{PCSK9}. The experimental maximal binding was obtained from the kinetic fitting and convert to %R_{max} so that the compound MW and PCSK9 density for each experiment can be normalized and compared.

Flow Cytometry

HepG2 cells were seeded at a density of 1.0x10^5 cells/well in 24-well plate with DMEM + 10% FBS medium for 18 h. After aspiration, cells were washed two times with PBS and treated with DMEM supplemented with 10% lipoprotein-deficient serum for another 18 h. Inhibitors (50 μM) or pep2-8 (50 μM) at 0.5 % DMSO level and 3 μg/mL PCSK9 in 10% lipoprotein-deficient DMEM were pre-incubated for 30 minutes at room temperature before adding into each well. Cells were incubated for 4 h at 37 ºC with 5% CO_2. After

treatment, cells were washed twice with PBS and detached with enzyme-free cell dissociation buffer (gibco) for 15 min at 37 ºC. After centrifugation (500 rcf, 5 mins, 4 ºC), cells were incubated with anti-LDLR antibody (100 µL of 10 µg/mL in 0.5% BSA/PBS, Invitrogen) on ice for 20 minutes. Cells were washed twice with 0.5% BSA/PBS and incubated with goat anti-mouse IgG (H+L) Alexa Fluor 647 (100 µL of 10 µg/mL in 0.5 % BSA/PBS, Invitrogen) on ice for 10 minutes. After washing twice with 0.5% BSA/PBS, cells were resuspended in 100 µL 0.5% BSA/PBS and analyzed by flow cytometer (BD Accuri™ C6 cytometer, BD Biosciences). The fluorescence of 30,000 events of each sample was acquired for data analysis. Forward scatter versus side scatter gates were set to exclude dead cells and debris. Mean values of the cells treated with PCSK9 were set as the minimum LDLR level (0%, negative), and without PCSK9 as the maximum LDLR level (100%, positive). Data from inhibitors were scaled as 0-100% according to the negative and positive controls. All experiments were repeated three times individually.

Fluorescent Imaging

HepG2 cells were seeded at the density of 1.0×10^4 cells/well in 24-well plate with DMEM + 10% FBS medium overnight, and then incubated in 10% lipoprotein-deficient DMEM for another 24 h. Inhibitors (50 µM) or Pep2-8 (50 µM) at 0.5 % DMSO level and 3 µg/mL PCSK9 in 10% lipoprotein-deficient DMEM were pre-incubated for 30 minutes at room temperature before adding into each well. Cells were incubated for 4 h at 37 ºC with 5% CO_2. After treatment, cells were washed twice with PBS and incubated with 300 µL anti-LDLR antibody (5 µg/mL in delipidated DMEM, Invitrogen) at 37 ºC for 30 minutes. Cells were washed twice with PBS and incubated with 300 µL goat anti-mouse IgG (H+L) Alexa Fluor 647 (10 µg/mL in delipidated DMEM, Invitrogen) at 37 ºC for 30

minutes. Cells were washed twice with PBS, and then treated with FluoroBrite™ DMEM (gibco), and then analyzed with fluorescent microscope (Evos FL Auto 2, Invitrogen). The average corrected total cell fluorescence (CTCF) was quantified by ImageJ software. Cells treated with PCSK9 alone were calibrated as 0% (negative), untreated with PCSK9 as 100% (positive), and cells treated with compounds and PCSK9 were scaled according to the controls.

Table D.2. *QikProp calculations of compounds **16***

Compounds	QPlogP$_{o/w}$	PSA (Å^3)	QPPCaco (nm/s)	Rule of Five	Rule of Three
DLLD-**16nclk**	-2.01	278	0.031	3	2
DLLD-**16nclr**	-2.59	319	0.023	3	2
DDDL-**16vclr**	-0.38	257	0.62	3	2
DDDD-**16vclr**	-0.34	262	0.55	3	2
LDLL-**16dlnr**	-5.68	366	0.001	3	2
LDLL-**16dl(CN)r**	-4.37	338	0.006	3	2
LLLL-**16qndw**	-6.19	362	0.001	3	2
LLLD-**16qndr**	-8.54	421	0	3	2
DLLL-**16qndr**	-8.54	421	0	3	2
LLDL-**16qndr**	-8.54	421	0	3	2
LLDD-**16qndr**	-8.54	421	0	3	2
DLDL-**16endr**	-7.12	416	0	3	2
DLDD-**16endh**	-6.45	368	0	3	2
LLLD-**16ldnq**	-7.08	348	0.002	3	2
LLLL-**16aaar**	-2.38	270	0.27	3	2

$QlogP_{o/w}$ – predicted octanol/water partition coefficient; PSA – Van der Waals surface area of polar nitrogen and oxygen atoms; QPPCaco – predicted Caco-2 cell permeability; Rule of Five – number of violations of Lipinski's rule of five; Rule of Three – number of violations of Jorgensen's rule of three.

*Solid-phase Syntheses and Characterizations of Compounds **16***

Loading methionine linker on resin: Tentagel-amine resin (600 mg, 0.26 mequiv/g) was swollen with CH_2Cl_2 (5 mL) in a fritted syringe for 15 min, and then in DMF for at least 1 h. After removing DMF, Fmoc-Met-OH (232 mg, 0.62 mmol), $^{i}Pr_2NEt$ (0.22 mL, 1.25 mmol) dissolved 2.5 mL of 0.25 M HATU/HOAt in DMF solution were added into the fritted syringe. The reaction was heated while stirring under microwave irradiation (100 W, 75 ºC, 10 min). Solution was drained, and washed with DMF (4 mL x 5). The completeness of loading was determined by negative Kaiser test.

Fmoc deprotection: The Fmoc deprotection was performed by shaking beads in 2 mL of 20% piperidine in DMF for 1 min, drained and the fresh 4 mL of 20% piperidine solution was added and heated under microwave irradiation (160 W, 75 ºC, 3 min). Solution was drained, and beads were washed with DMF (4 mL x 5), CH_2Cl_2 (4 mL x 3), MeOH (4 mL x 3) and DMF (4 mL x 3).

Amide coupling: Fmoc-amino acids or nosyl-containing pip-acids (4 eq) and $^{i}Pr_2NEt$ (8 eq) dissolved 2.5 mL of 0.25 M HATU/HOAt in DMF solution were added into the resin containing primary- or secondary-amine. Beads were heated under microwave irradiation (100 W, 75 ºC, 10 min), then washed with 4 mL x 5 of DMF. For arginine amino acid, beads were shaken in the coupling solution for 20 minutes at room

temperature, and then irradiated with microwave (100 W, 75 ºC, 10 min). The solution was changed with the fresh Fmoc-arginine coupling solution and irradiated again with the same condition. For cysteine and histidine amino acids, the coupling reaction was performed under microwave at lower temperature to avoid epimerization (100 W, 50 ºC, 15 min).

Nosyl deprotection: After coupling with nosyl-containing pip-acids, beads were shaken in the solution of mercaptoethanol (5 eq) and DBU (5 eq) in 3 mL of DMF for 15 min at room temperature. The orange solution was drained, washed thrice with 3 mL DMF, and beads were shaken again with the fresh solution of mercaptoethanol/DBU for 15 min at room temperature. Beads were washed 5 times with DMF, and a few beads were analyzed by chloranil test to confirm the presence of secondary amine.

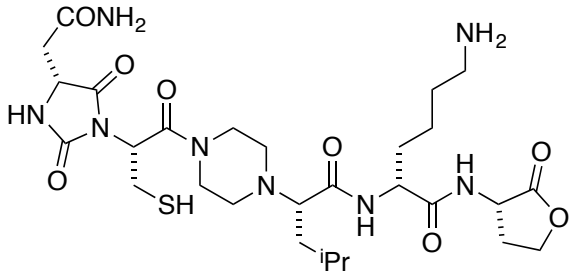
Hydantoin cyclization: After the Fmoc deprotection at the R^1 position, beads were treated twice with the solution of 2 equiv of *p*-nitrophenyl chloroformate and 4 equiv of $^{i}Pr_2NEt$ in MeCN at ~ 0.05 M concentration. Beads were shaken at room temperature for 30 min each. After coupling step, beads were washed with MeCN (4 mL x 3), CH_2Cl_2 (4 mL x 3) and NMP (4 mL x 3). Beads were shaken 5 times with 5 mL of 5 % $^{i}Pr_2NEt$ /NMP solution for 1 h each to remove any byproducts. Beads were washed with NMP (4 mL x 5), CH_2Cl_2 (4 mL x 5) and MeOH (4 mL x 5) and dried.

Cleavage from the resin: Compounds were cleaved off from the beads by treating with cyanogen bromide (30 mg/mL) in MeCN/acetic acid/H_2O (5:4:1 v/v) solution. Beads were shaken at room temperature for at least 24 h. After filtration, the crude material was dried under nitrogen stream and purified by preparative reversed-phase HPLC (50%

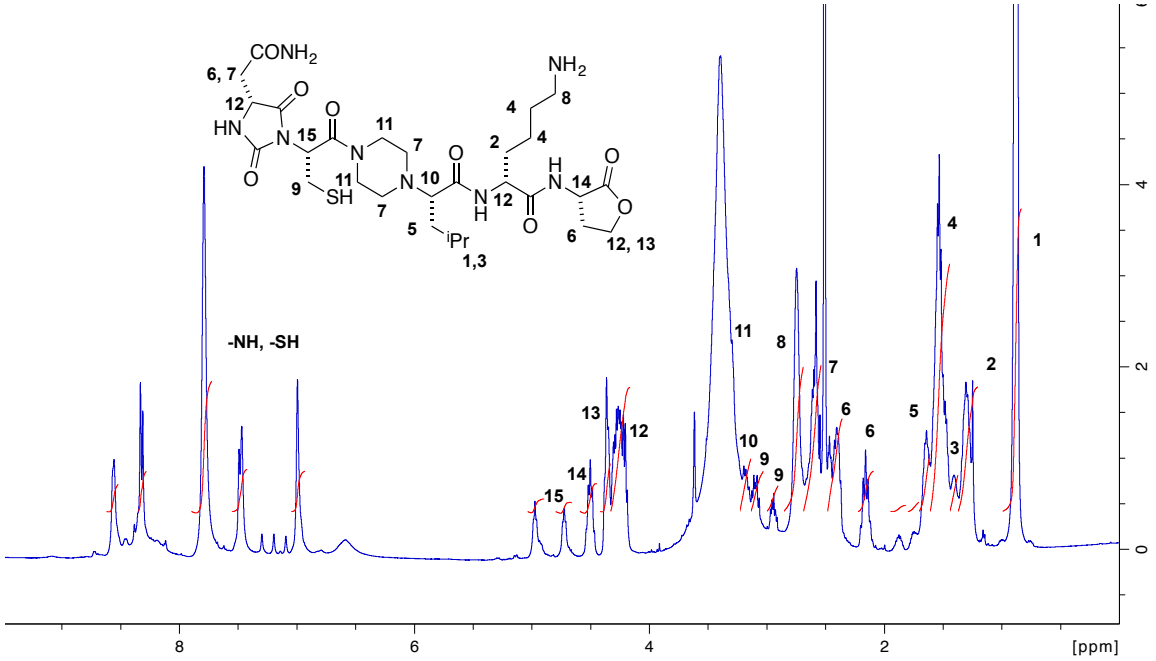
- 95% MeCN/water containing 0.05% TFA). Products were lyophilized to obtain white powders.

Side-chain deprotection: Purified compounds were stirred in TFA/Et_3SiH/water (95:2.5:2.5 v/v) cocktail for 4 h at room temperature. Crude materials were dried under vacuum, and then precipitated with cold *tert*-butyl methyl ether. Solid compounds were collected, washed several times with *tert*-butyl methyl ether to remove any byproducts. Compounds were dissolved in water and lyophilized to obtain final products as white solids.

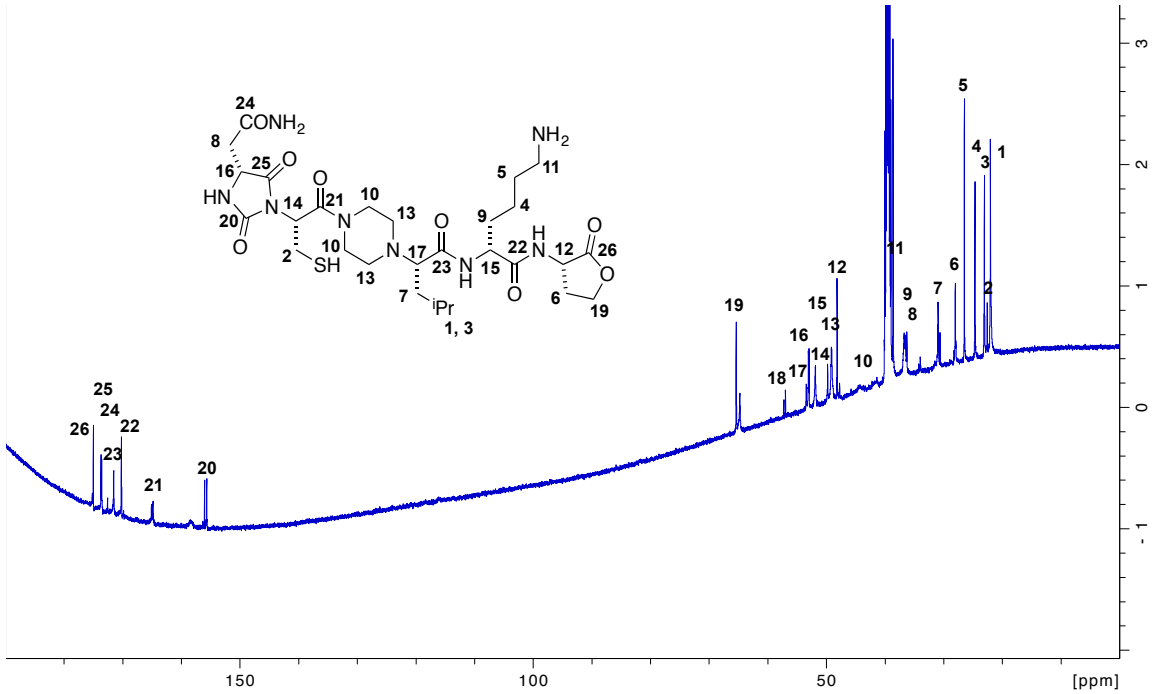
(*R*)-6-Amino-2-((*S*)-2-(4-((*R*)-2-((*R*)-4-(2-amino-2-oxoethyl)-2,5-dioxoimidazolidin-1-yl)-3-mercaptopropanoyl)piperazin-1-yl)-4-methylpentanamido)-*N*-((*S*)-2-oxotetrahydrofuran-3-yl)hexanamide (DLLD-**16nclk**)



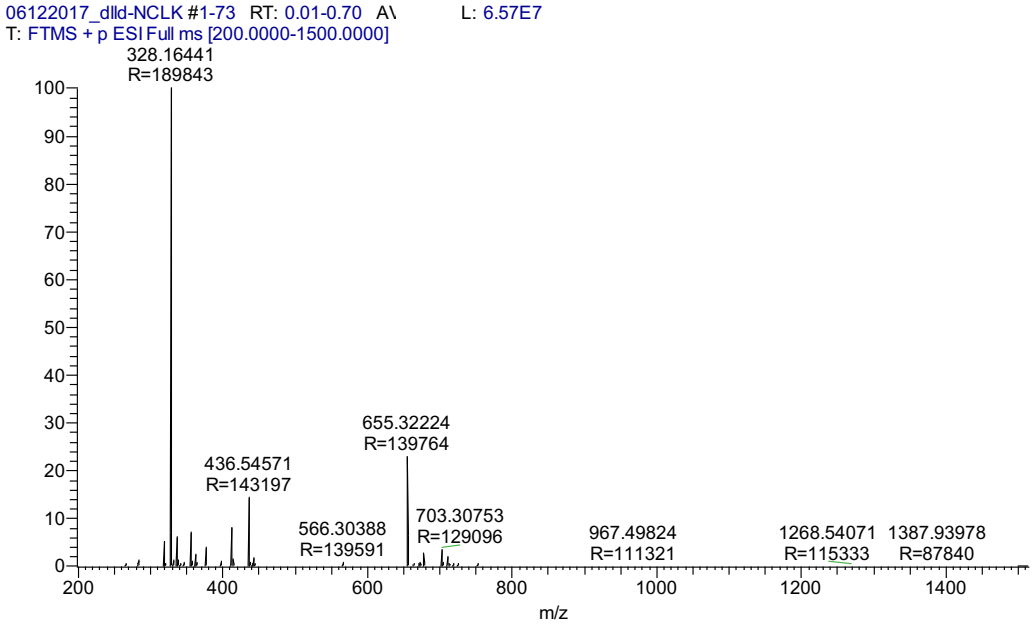
Data for DLLD-16nclk. ^{1}H NMR (400 MHz, DMSO-d_6) δ 8.56 (br, 1H), 8.32 (m, 1H), 7.79 (m, 4H), 7.48 (m, 1H), 6.99 (m, 1H), 4.72 (m, 1H), 4.50 (m, 1H), 4.37 (m, 1H), 4.33-4.17 (m, 3H), 3.49-3.39 (m, 4H), 3.37 (m, 1H), 3.10 (m, 1H), 2.94 (m, 1H), 2.75 (m, 2H), 2.64-2.54 (m, 5H), 2.49-2.36 (m, 2H), 2.16 (m, 1H), 1.69-1.22 (m, 9H), 0.88 (m, 6H);



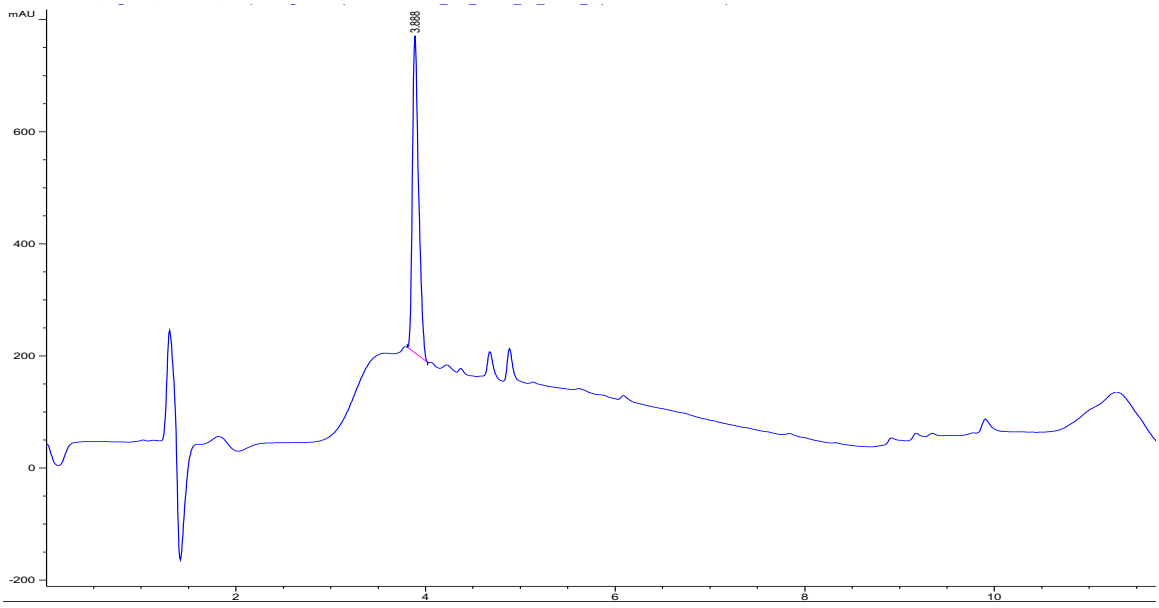
^{13}C NMR (100 MHz, DMSO-d_6) δ 175.0, 173.7, 173.6, 171.5, 170.2, 165.0, 155.7, 65.4, 57.1, 53.4, 53.1, 52.9, 51.9, 49.8, 49.1, 48.2, 44.5, 41.8, 38.6, 36.5, 36.3, 31.0, 28.0, 26.4, 24.7, 23.0, 22.6, 22.0



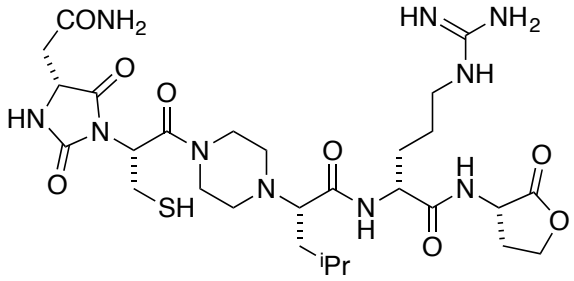
HRMS (ESI) m/z calcd for $C_{28}H_{47}N_8O_8S^+$ 655.3232; found 655.3222 $(M+H)^+$



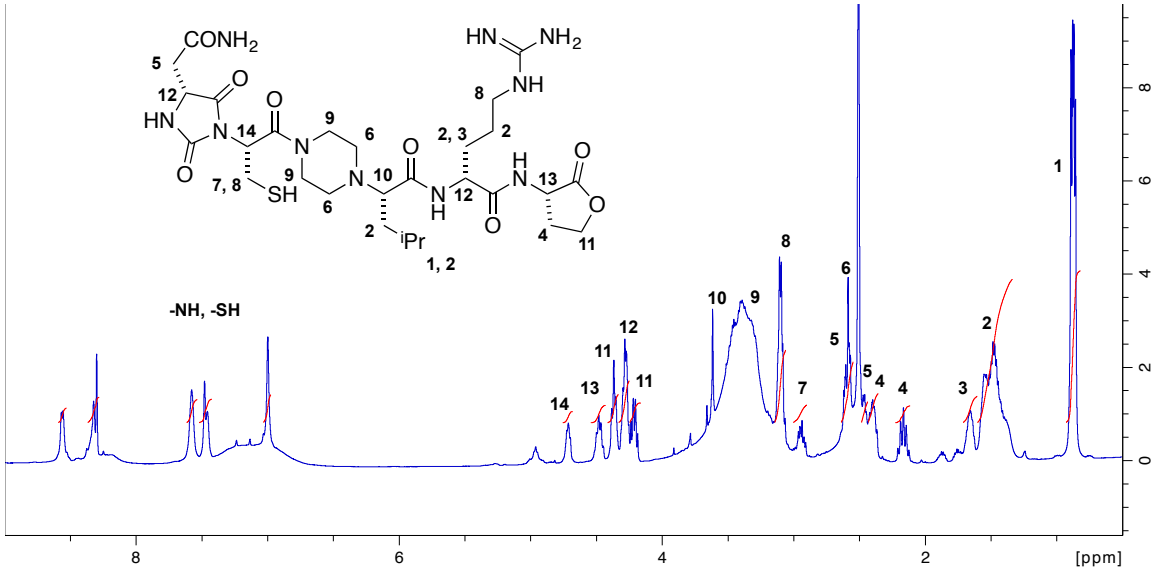
LC



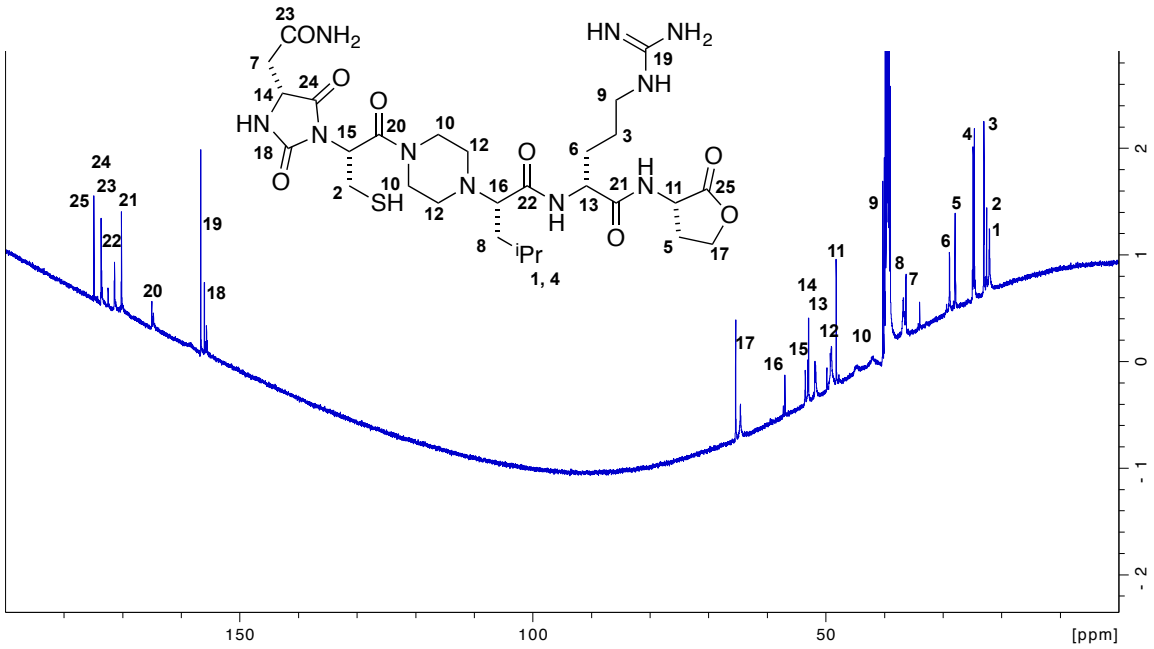
(*S*)-2-(4-((*R*)-2-((*R*)-4-(2-Amino-2-oxoethyl)-2,5-dioxoimidazolidin-1-yl)-3-mercaptopropanoyl)piperazin-1-yl)-*N*-((*R*)-5-guanidino-1-oxo-1-(((*S*)-2-oxotetrahydrofuran-3-yl)amino)pentan-2-yl)-4-methylpentanamide (DLLD-**16nclr**)



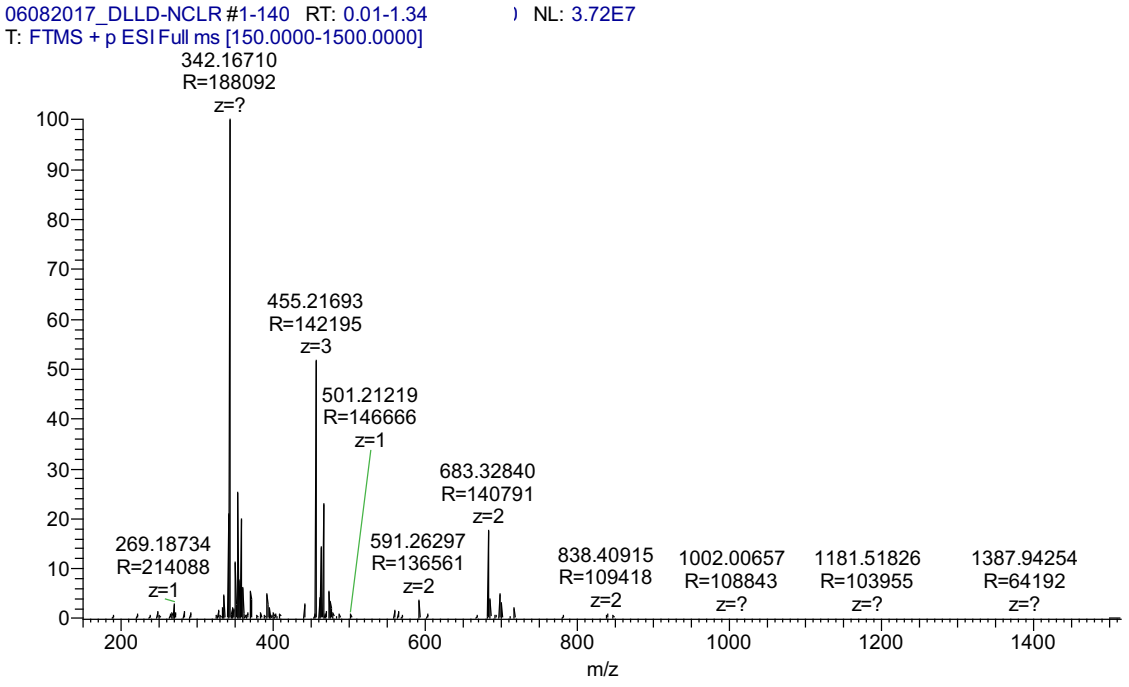
Data for DLLD-16nclr. ^{1}H NMR (400 MHz, DMSO-d_6) δ 8.57 (m, 1H), 8.32 (m, 1H), 7.58 (br, 1H), 7.47 (m, 1H), 7.00 (br, 1H), 4.72 (m, 1H), 4.48 (m, 1H), 4.37 (m, 1H), 4.28 (m, 2H), 4.21 (m, 1H), 3.62 (m, 1H), 3.40-3.23 (m, 4H), 3.10 (m, 3H), 2.94 (m, 1H), 2.63-2.55 (m, 3H), 2.53-2.48 (m, 2H), 2.48-2.44 (m, 1H), 2.44-2.36 (m, 1H), 2.17 (m, 1H), 1.66 (m, 1H), 1.61-1.32 (m, 6H), 0.88 (m, 6H);



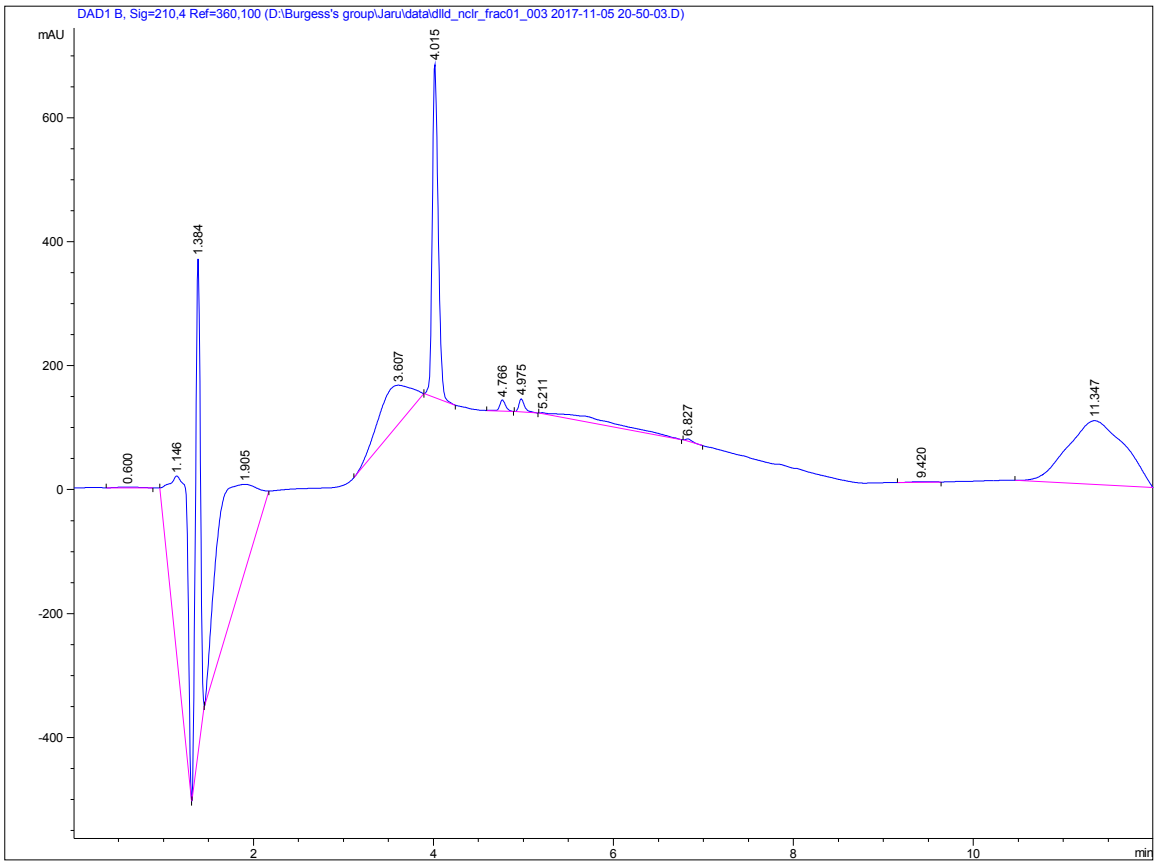
^{13}C NMR (100 MHz, DMSO-d_6) δ 174.9, 173.7, 172.5, 171.4, 170.2, 165.0, 156.7, 156.0, 65.4, 64.6, 57.0, 53.5, 52.9, 51.8, 49.0, 48.3, 44.8, 42.1, 40.4, 36.8, 36.4, 28.9, 28.0, 24.9, 23.0, 22.5, 22.1



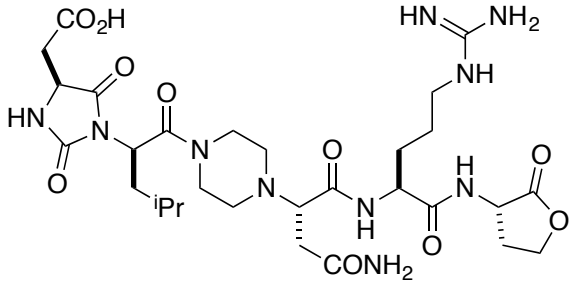
HRMS (ESI) m/z calcd for $C_{28}H_{47}N_{10}O_8S^+$ 683.3294; found 683.3284 $(M+H)^+$



LC

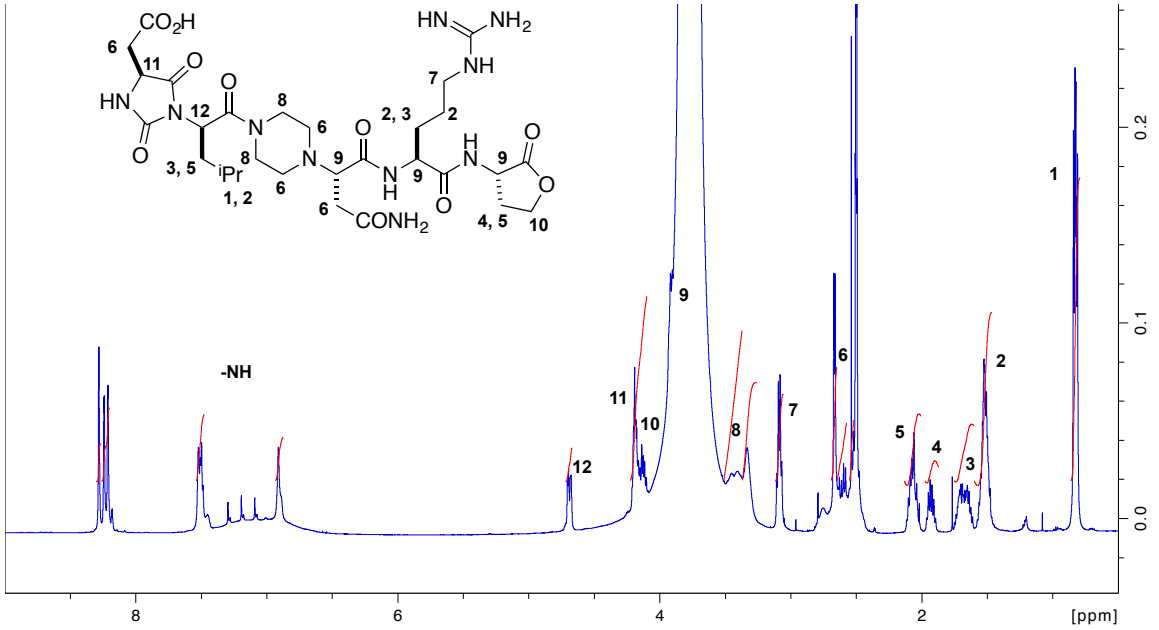


2-((*S*)-1-((*R*)-1-(4-((*S*)-4-Amino-1-(((*S*)-5-guanidino-1-oxo-1-(((*S*)-2-oxotetrahydrofuran-3-yl)amino)pentan-2-yl)amino)-1,4-dioxobutan-2-yl)piperazin-1-yl)-4-methyl-1-oxopentan-2-yl)-2,5-dioxoimidazolidin-4-yl)acetic acid (LDLL-**16dlnr**)

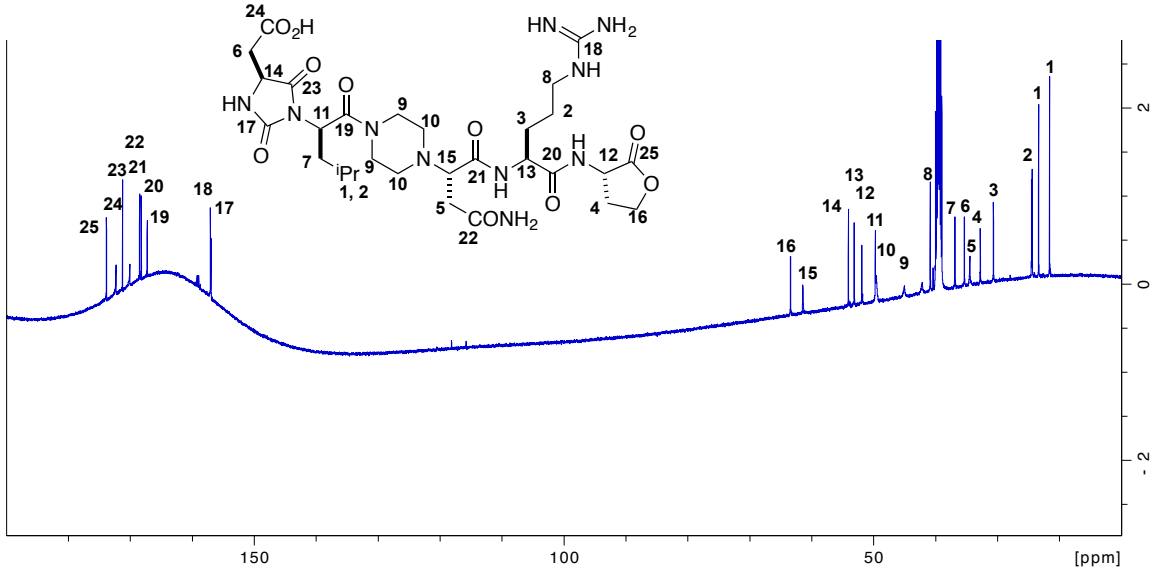


Data for LDLL-16dlnr. ^{1}H NMR (400 MHz, DMSO-d_6) δ 8.31-8.18 (m, 2H), 7.57-7.44 (m, 1H), 6.92 (br, 1H), 4.70 (m, 1H), 4.20 (m, 1H), 4.13 (m, 1H), 3.91 (m, 1H), 3.86 (m, 1H),

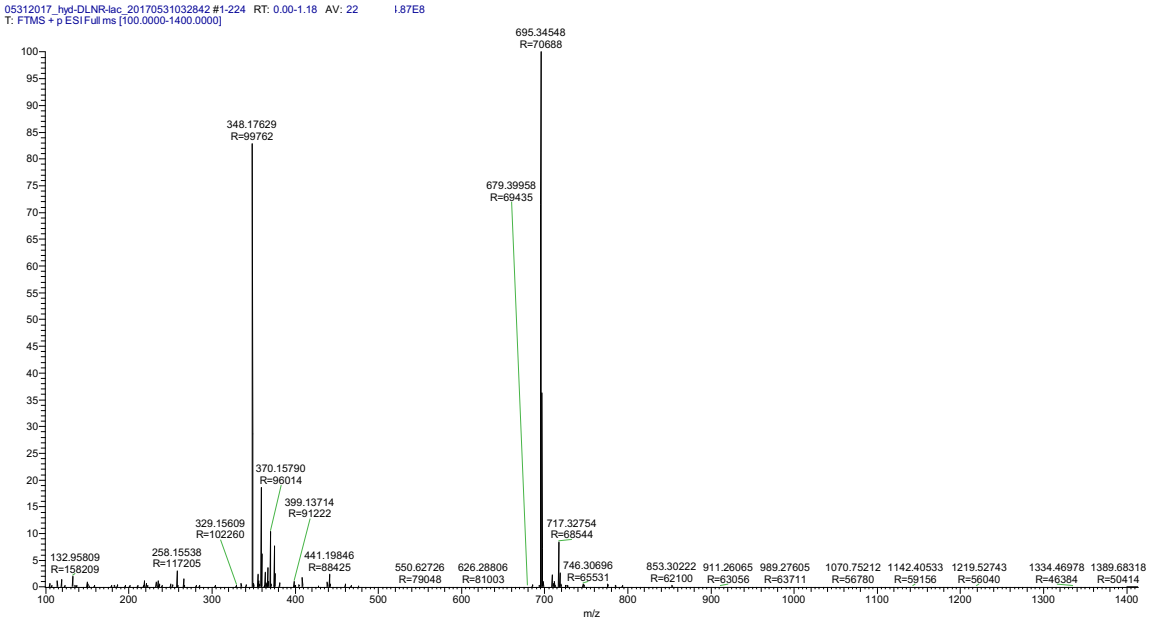
3.76 (m, 1H), 3.51-3.25 (m, 4H), 3.10 (m, 2H), 2.68 (d, *J* = 4.9 Hz, 2H), 2.65-2.45 (m, 6H), 2.08 (m, 2H), 1.94 (m, 1H), 1.69 (m, 2H), 1.53 (m, 6H), 0.84 (dd, *J* = 9.4, 6.2 Hz, 6H)



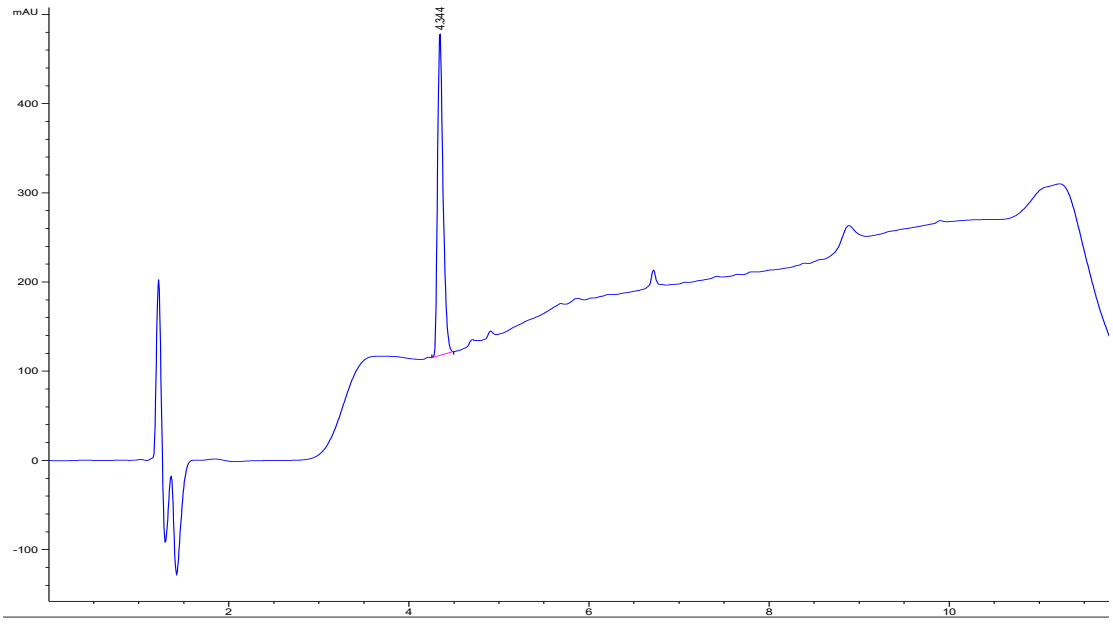
^{13}C NMR (100 MHz, DMSO-d_6) δ; 173.9, 172.3, 171.3, 170.1, 168.5, 168.3, 167.3, 157.1, 157.0, 63.5, 61.5, 54.1, 53.2, 51.9, 49.8, 49.6, 45.1, 42.2, 40.9, 36.9, 35.4, 34.5, 32.8, 30.7, 24.6, 24.4, 23.4, 21.6



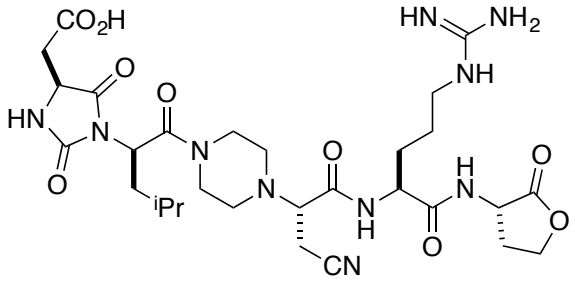
HRMS (ESI) m/z calcd for $C_{29}H_{47}N_{10}O_{10}^{+}$ 695.3471; found 695.3455 $(M+H)^{+}$



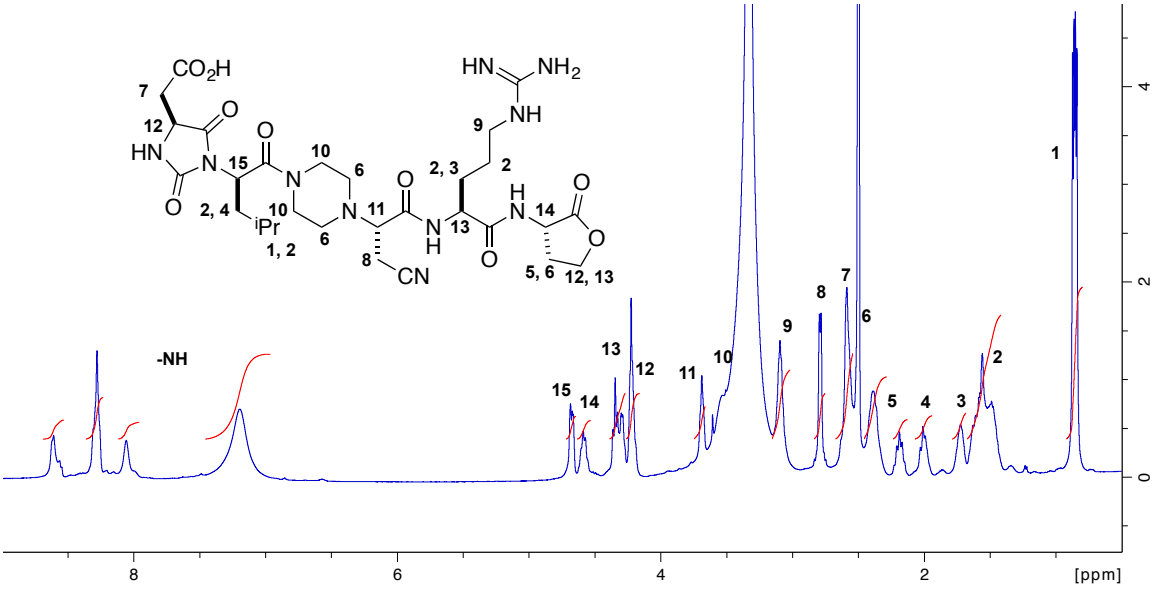
LC



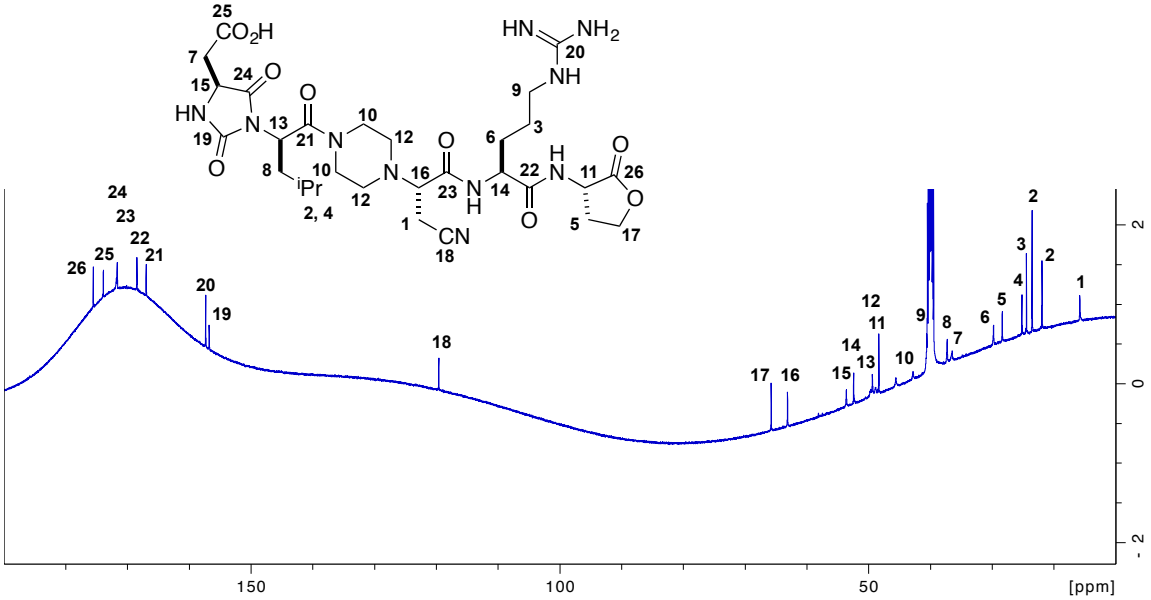
2-((*S*)-1-((*R*)-1-(4-((*S*)-3-Cyano-1-(((*S*)-5-guanidino-1-oxo-1-(((*S*)-2-oxotetrahydrofuran-3-yl)amino)pentan-2-yl)amino)-1-oxopropan-2-yl)piperazin-1-yl)-4-methyl-1-oxopentan-2-yl)-2,5-dioxoimidazolidin-4-yl)acetic acid (LDLL-**16dl(CN)r**)



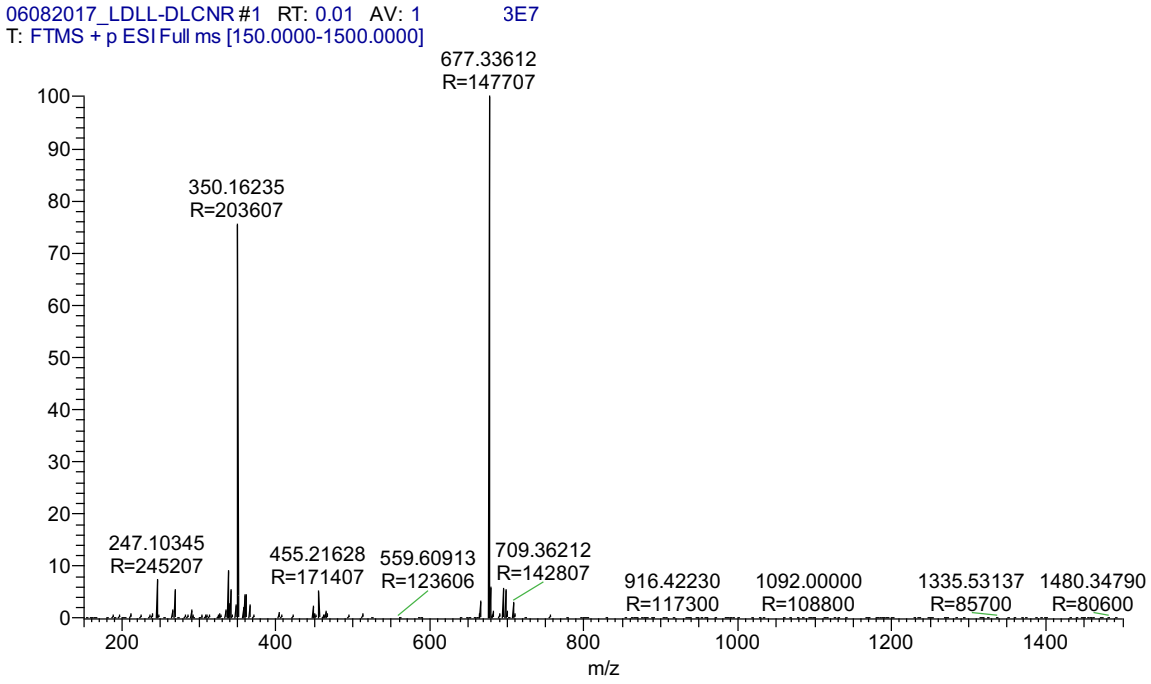
Data for LDLL-16dl(CN)r. ^{1}H NMR (400 MHz, DMSO-d_6) δ 8.61 (br, 1H), 8.34-8.23 (br, 2H), 8.06 (br, 1H), 7.45-6.97 (br, 3H), 4.68 (m, 1H), 4.59 (m, 1H), 4.38-4.26 (m, 2H), 4.26-4.16 (m, 2H), 3.69 (m, 1H), 3.59-3.45 (m, 4H), 3.10 (m, 2H), 2.79 (m, 2H), 2.67-2.54 (m, 4H), 2.46-2.29 (m, 3H), 2.19 (m, 1H), 2.01 (m, 1H), 1.72 (m, 1H), 1.67-1.41 (m, 5H), 0.87 (d, *J* = 6.1 Hz, 3H), 0.85 (d, *J* = 6.4 Hz, 3H)



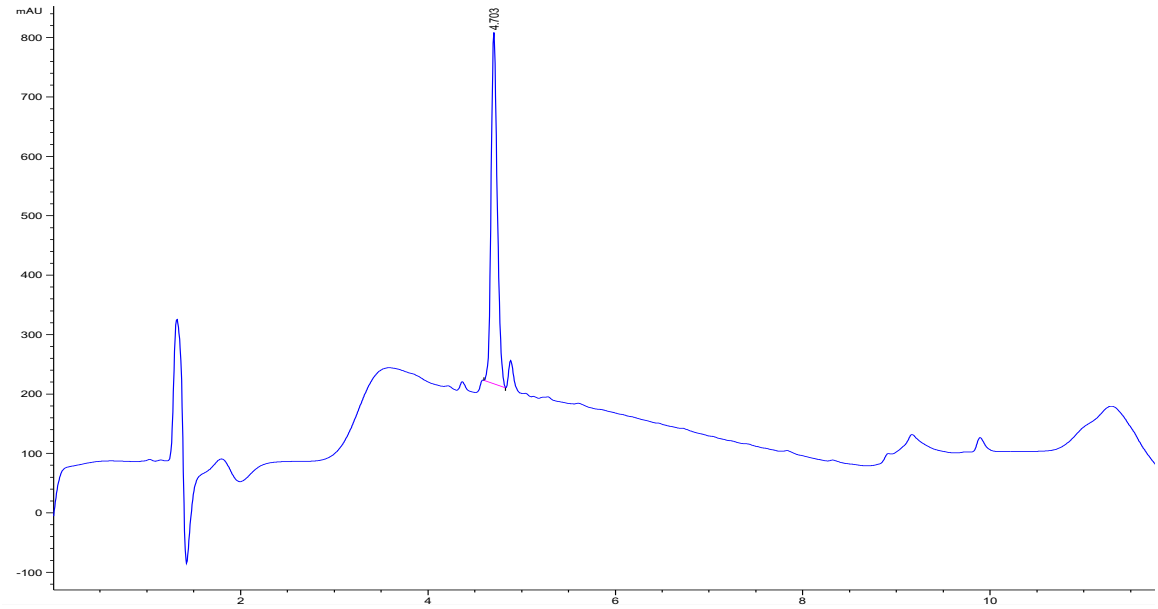
^{13}C NMR (100 MHz, DMSO-d_6) δ 175.0, 173.4, 171.3, 171.2, 168.0, 166.5, 156.8, 156.3, 119.1, 65.3, 62.7, 53.1, 51.9, 49.2, 48.9, 48.4, 47.9, 45.1, 42.3, 40.2, 36.8, 36.0, 29.2, 27.9, 24.7, 24.0, 23.0, 21.4, 15.3



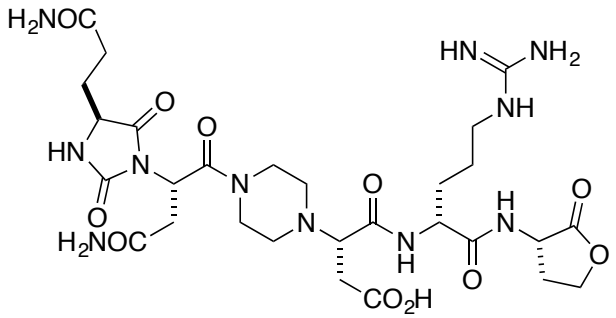
HRMS (ESI) m/z calcd for $C_{29}H_{45}N_{10}O_9^+$ 677.3365; found 677.3361 $(M+H)^+$



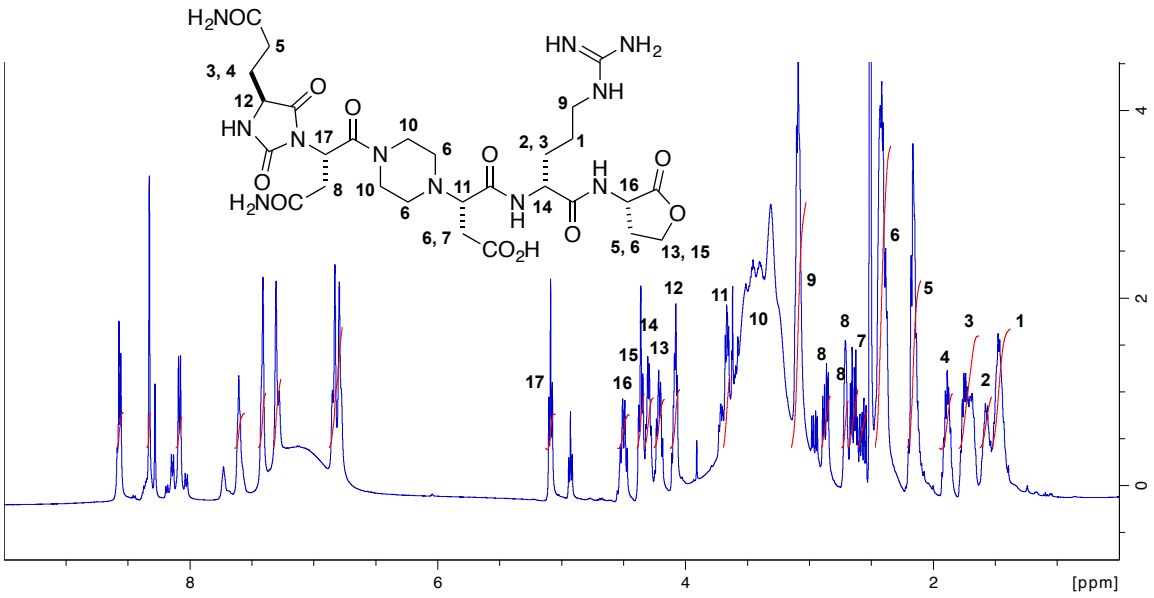
LC



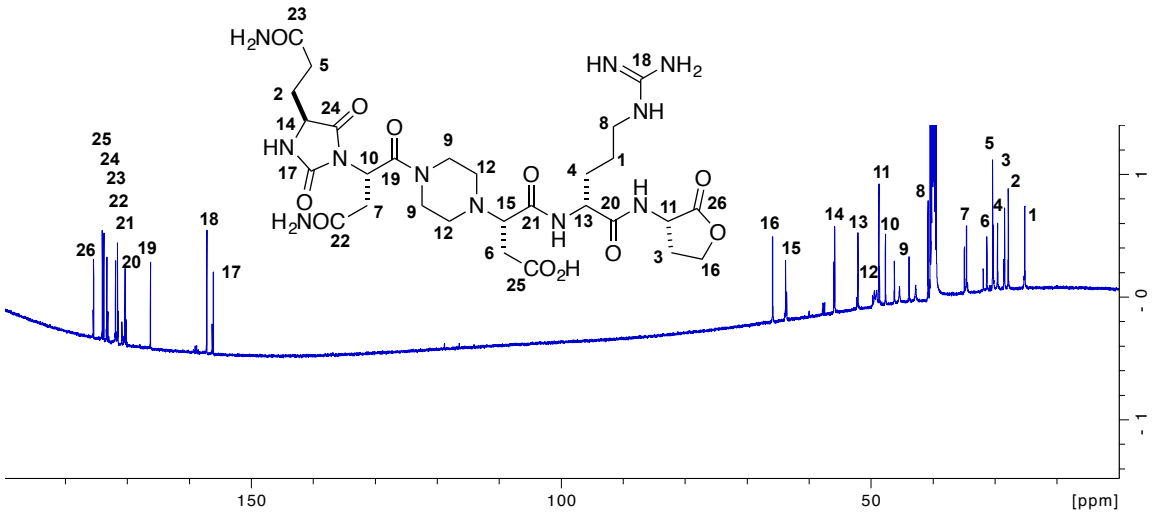
(*S*)-3-(4-((*S*)-4-Amino-2-((*S*)-4-(3-amino-3-oxopropyl)-2,5-dioxoimidazolidin-1-yl)-4-oxobutanoyl)piperazin-1-yl)-4-(((*R*)-5-guanidino-1-oxo-1-(((*S*)-2-oxotetrahydrofuran-3-yl)amino)pentan-2-yl)amino)-4-oxobutanoic acid (LLLD-**16qndr**)



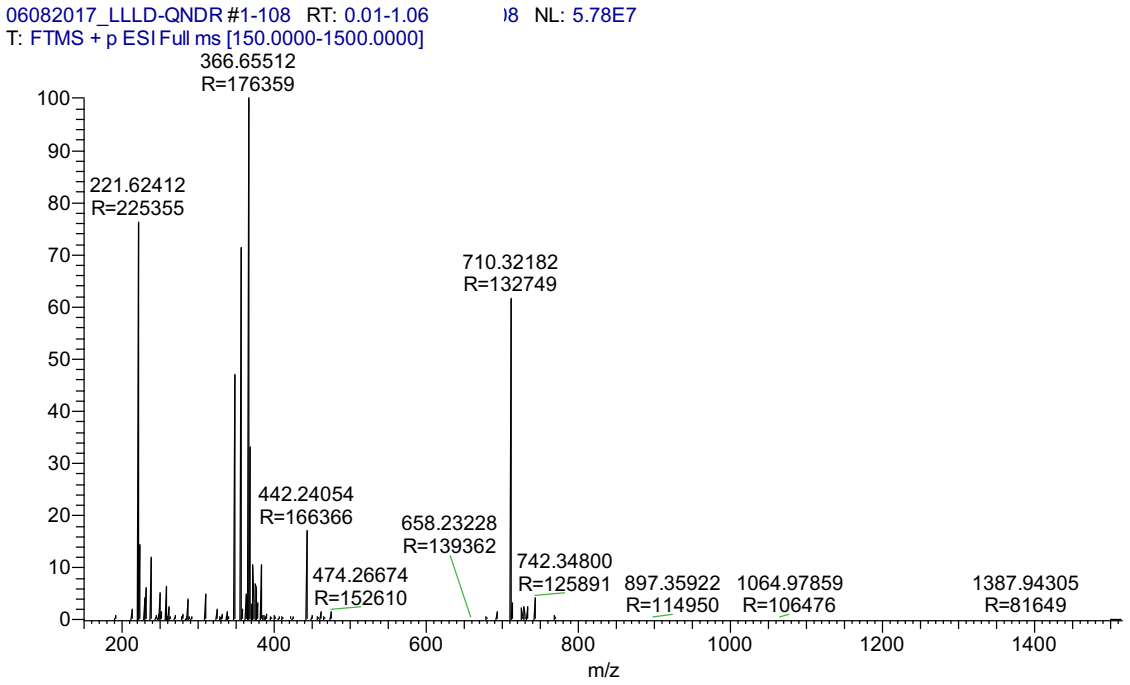
Data for LLLD-16qndr. ^{1}H NMR (400 MHz, DMSO-d_6) δ 8.57 (m, 1H), 8.33 (s, 1H), 8.25 (d, *J* = 8.3 Hz, 1H), 7.60 (br, 1H), 7.41 (br, 2H), 7.30 (br, 2H), 6.81 (m, 3H), 5.09 (t, *J* = 7.0 Hz, 1H), 4.50 (m, 1H), 4.36 (m, 1H), 4.30 (m, 1H), 4.21 (m, 1H), 4.09 (m, 1H), 3.67 (m, 1H), 3.42-3.17 (m, 4H), 3.09 (m, 2H), 2.87 (dd, *J* = 15.5, 7.1 Hz, 1H), 2.65 (dd, *J* = 15.3, 6.8 Hz, 1H), 2.56 (m, 1H), 2.42 (m, 6H), 2.16 (m, 3H), 1.89 (m, 1H), 1.75 (m, 1H), 1.69 (m, 1H), 1.58 (m, 1H), 1.47 (m, 2H)



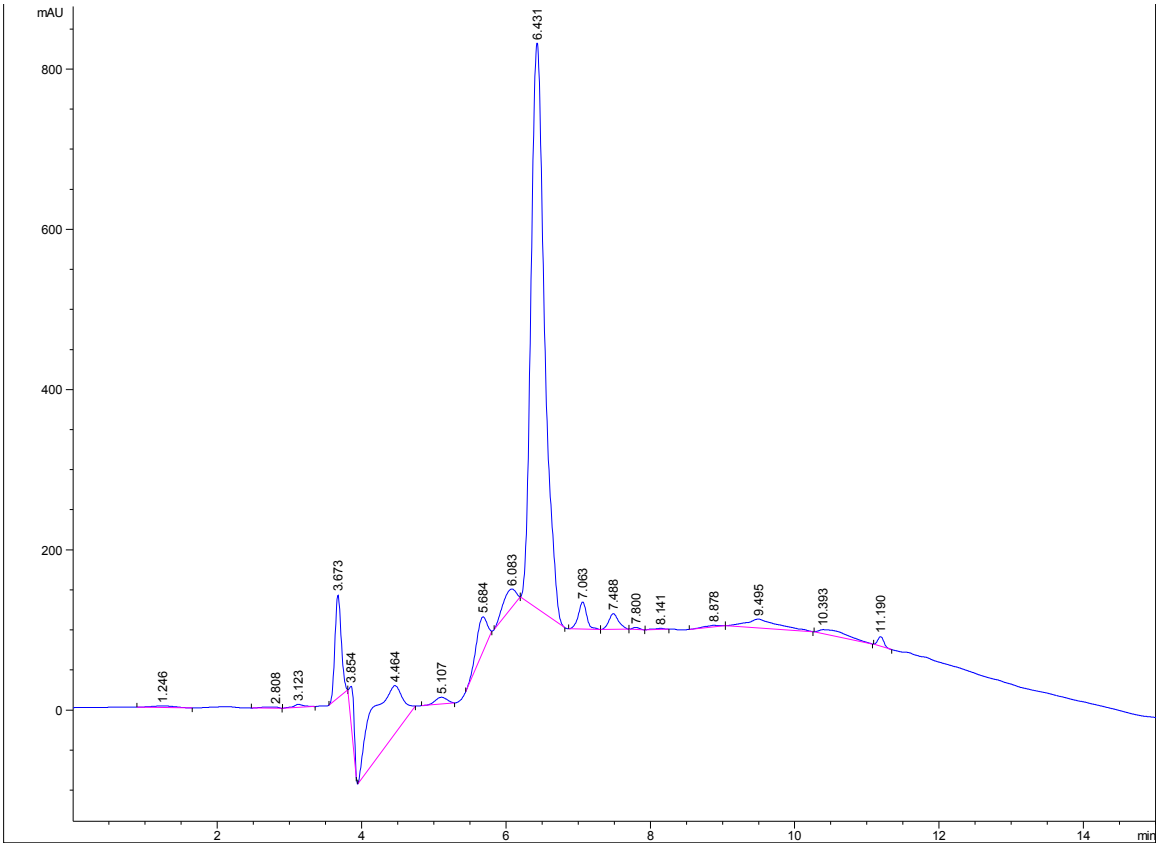
^{13}C NMR (100 MHz, DMSO-d_6) δ 175.0, 173.6, 173.3, 172.8, 171.4, 171.1, 169.9, 165.8, 156.7, 155.7, 65.4, 63.4, 55.4, 51.7, 48.9, 48.3, 47.2, 45.0, 42.3, 40.4, 34.2, 30.9, 30.0, 29.1, 28.0, 27.4, 24.7



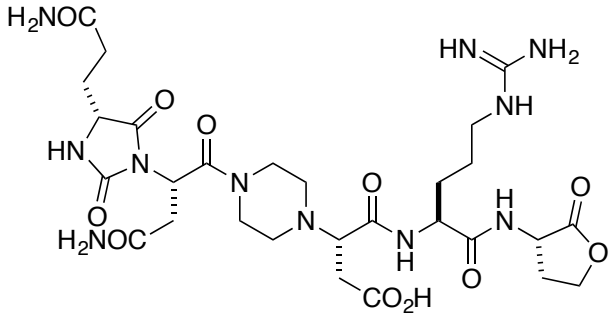
HRMS (ESI) m/z calcd for $C_{28}H_{44}N_{11}O_{11}^{+}$ 710.3216; found 710.3218 $(M+H)^{+}$



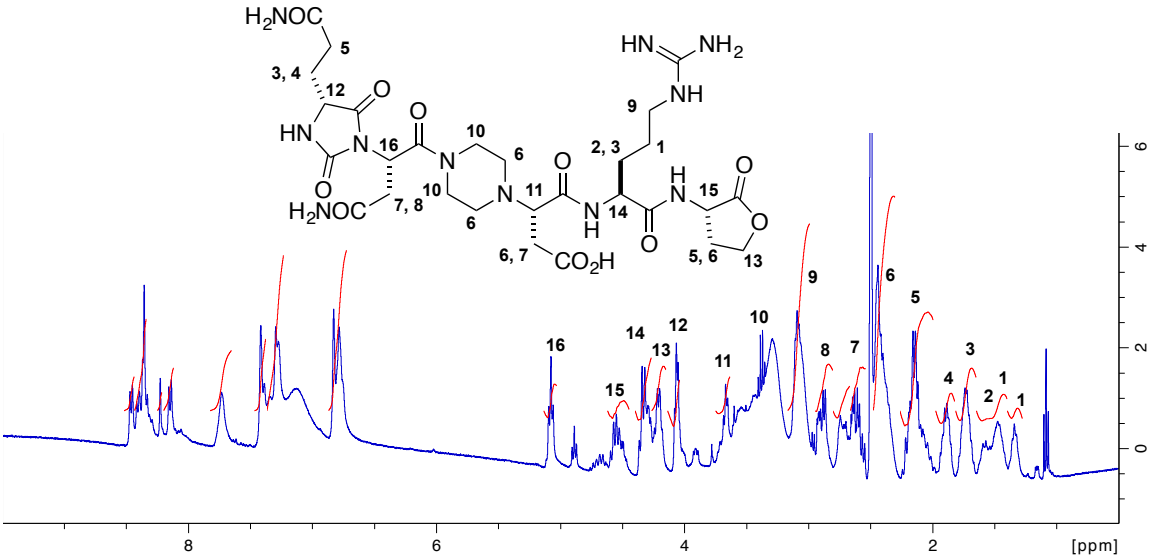
LC



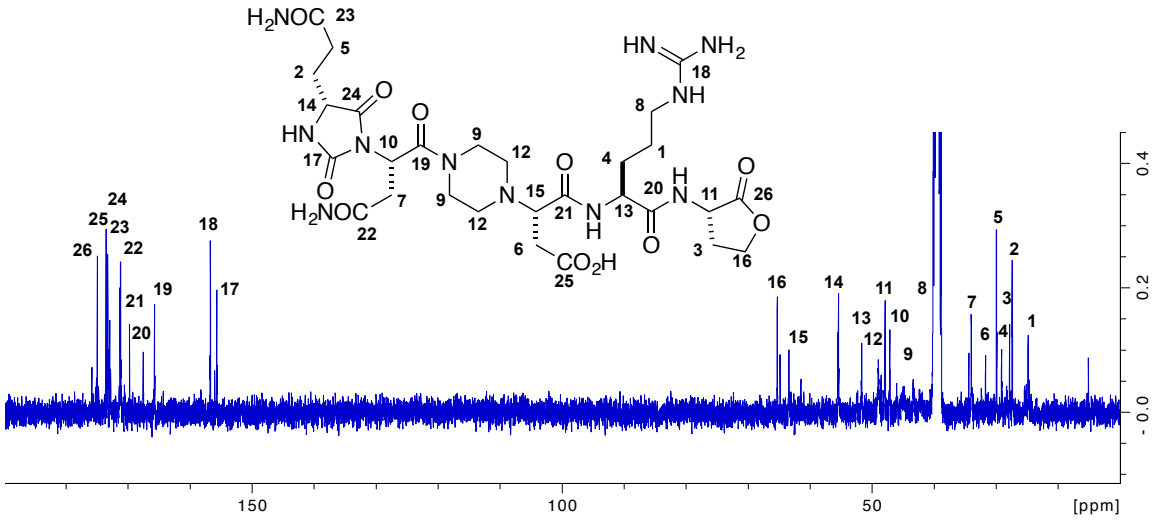
(*S*)-3-(4-((*S*)-4-Amino-2-((*R*)-4-(3-amino-3-oxopropyl)-2,5-dioxoimidazolidin-1-yl)-4-oxobutanoyl)piperazin-1-yl)-4-(((*S*)-5-guanidino-1-oxo-1-(((*S*)-2-oxotetrahydrofuran-3-yl)amino)pentan-2-yl)amino)-4-oxobutanoic acid (DLLL-**16qndr**)



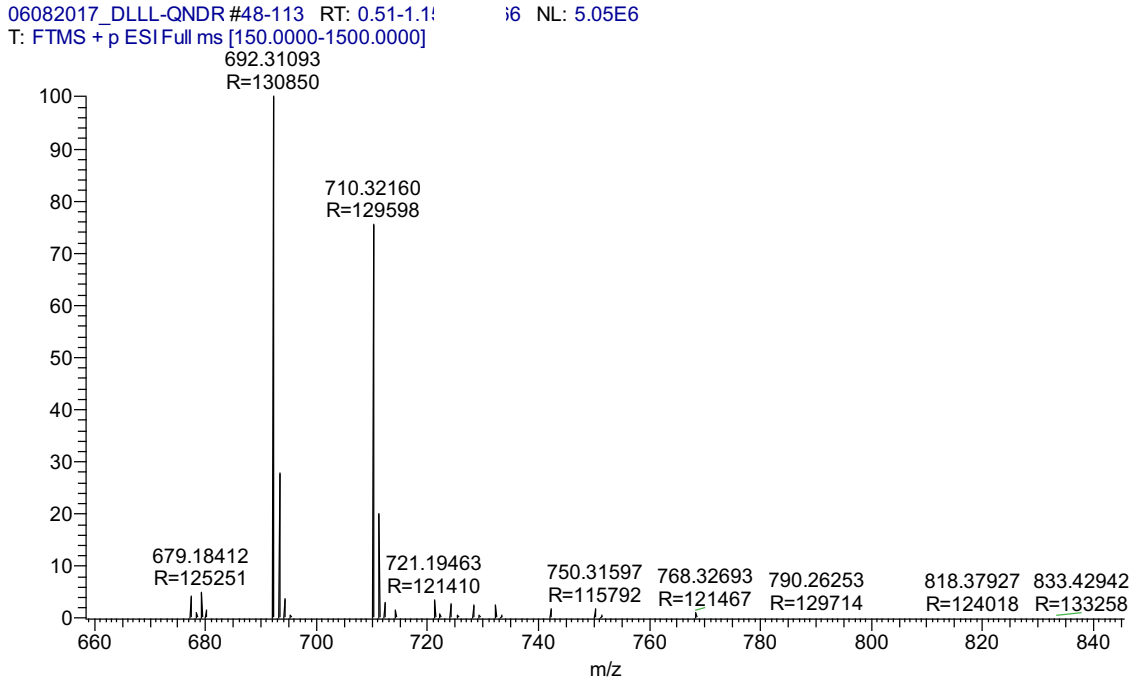
Data for DLLL-16qndr. ^{1}H NMR (400 MHz, DMSO-d$_6$) δ 8.46 (d, *J* = 7.7 Hz, 1H), 8.43-8.34 (m, 3H), 8.23 (br, 1H), 8.15 (d, *J* = 8.2 Hz, 1H), 7.73 (br, 1H), 7.42 (m, 1H), 7.28 (m, 2H), 6.81 (m, 2H), 5.08 (m, 1H), 4.53 (m, 1H), 4.39-4.26 (m, 2H), 4.21 (m, 1H), 4.07 (m, 1H), 3.67 (m, 1H), 3.29 (m, 4H), 3.08 (m, 2H), 2.89 (m, 1H), 2.67-2.53 (m, 2H), 2.48-2.30 (m, 6H), 2.26-1.99 (m, 3H), 1.90 (m, 1H), 1.73 (m, 2H), 1.64-1.40 (m, 2H), 1.34 (m, 1H)



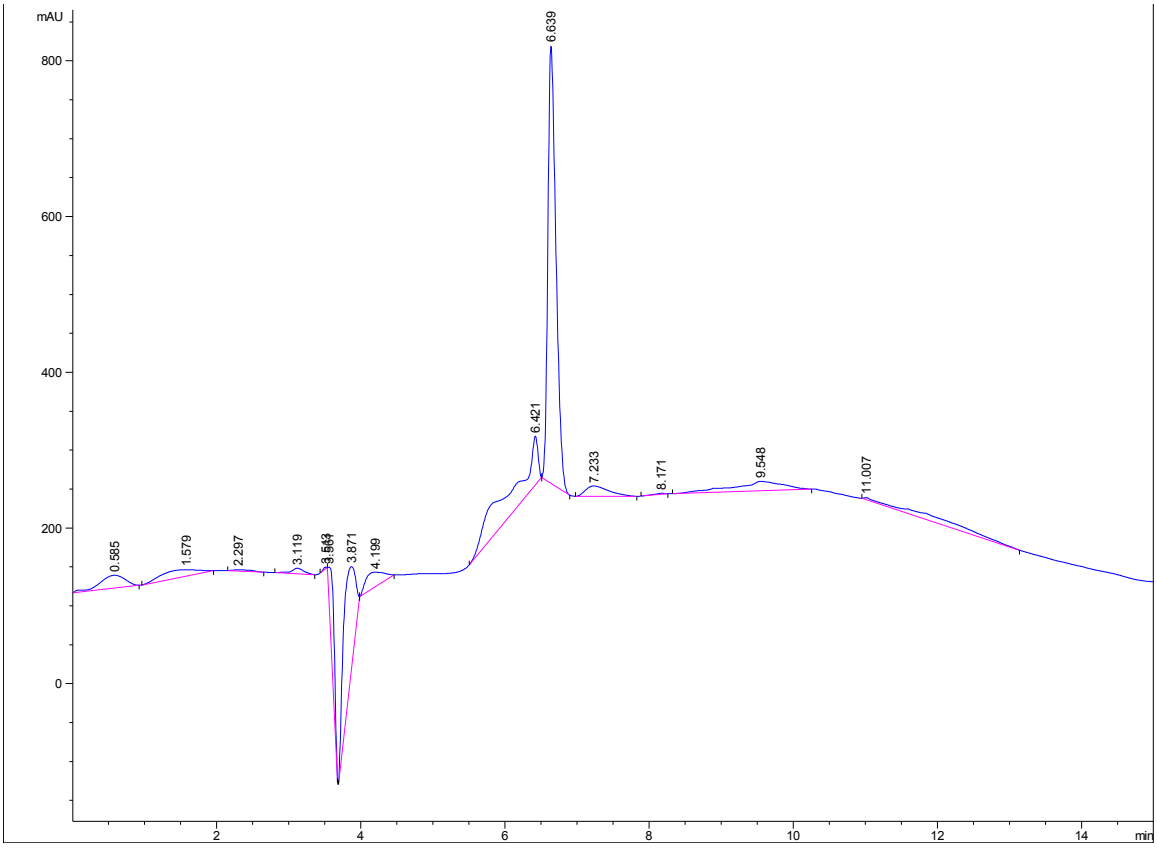
^{13}C NMR (100 MHz, DMSO-d_6) δ 175.0, 173.5, 173.3, 171.3, 171.2, 169.8, 167.5, 165.7, 156.7, 155.7, 65.3, 63.5, 55.4, 51.8, 49.0, 48.0, 47.2, 45.1, 42.2, 40.4, 34.1, 31.7, 30.0, 29.2, 27.8, 27.5, 24.9



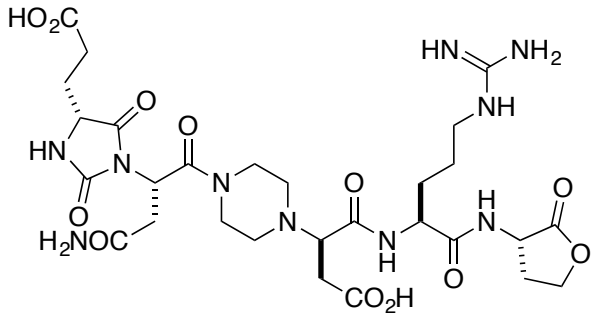
HRMS (ESI) m/z calcd for $C_{28}H_{44}N_{11}O_{11}^+$ 710.3216; found 710.3216 $(M+H)^+$



LC

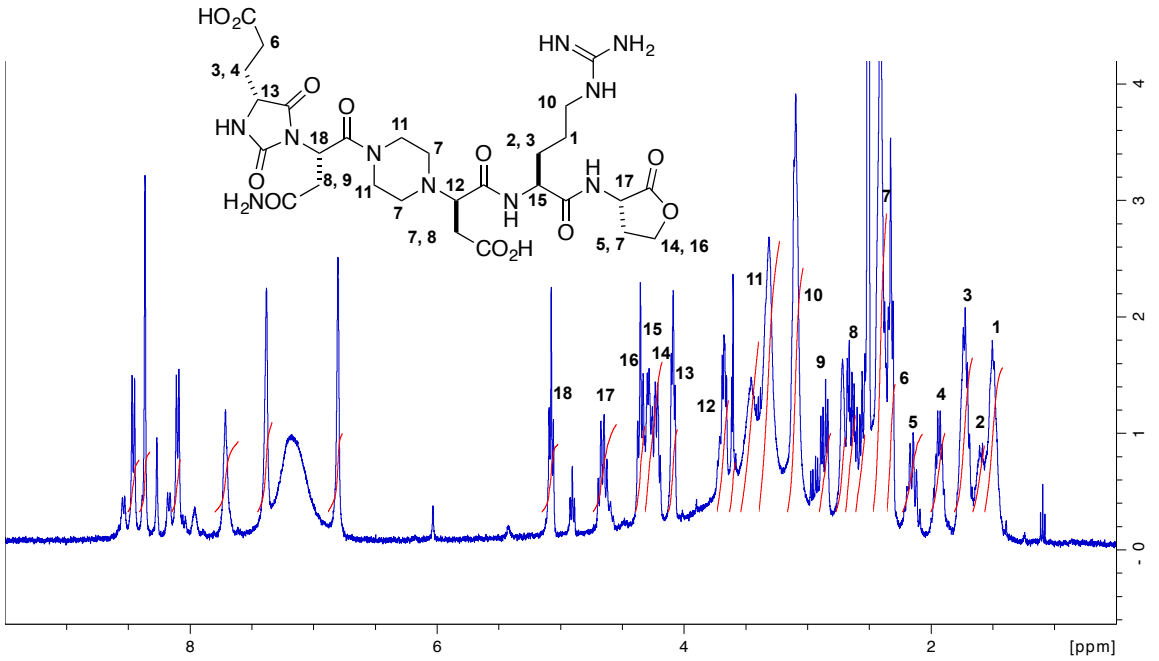


(*R*)-3-(4-((*S*)-4-Amino-2-((*R*)-4-(2-carboxyethyl)-2,5-dioxoimidazolidin-1-yl)-4-oxobutanoyl)piperazin-1-yl)-4-(((*S*)-5-guanidino-1-oxo-1-(((*S*)-2-oxotetrahydrofuran-3-yl)amino)pentan-2-yl)amino)-4-oxobutanoic acid (DLDL-**16endr**)

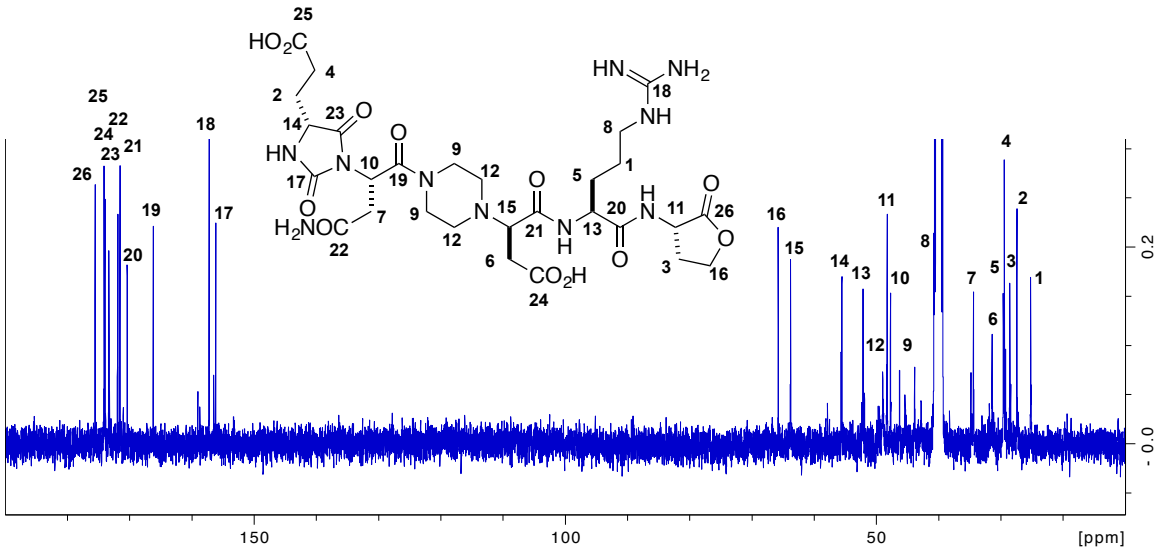


Data for DLDL-16endr. ^{1}H NMR (400 MHz, DMSO-d_6) δ 8.46 (d, *J* = 8.0 Hz, 1H), 8.37 (s, 1H), 8.10 (d, *J* = 8.3 Hz, 1H), 7.71 (br, 1H), 7.38 (br, 1H), 7.31-6.97 (br, 4H), 6.80 (br,

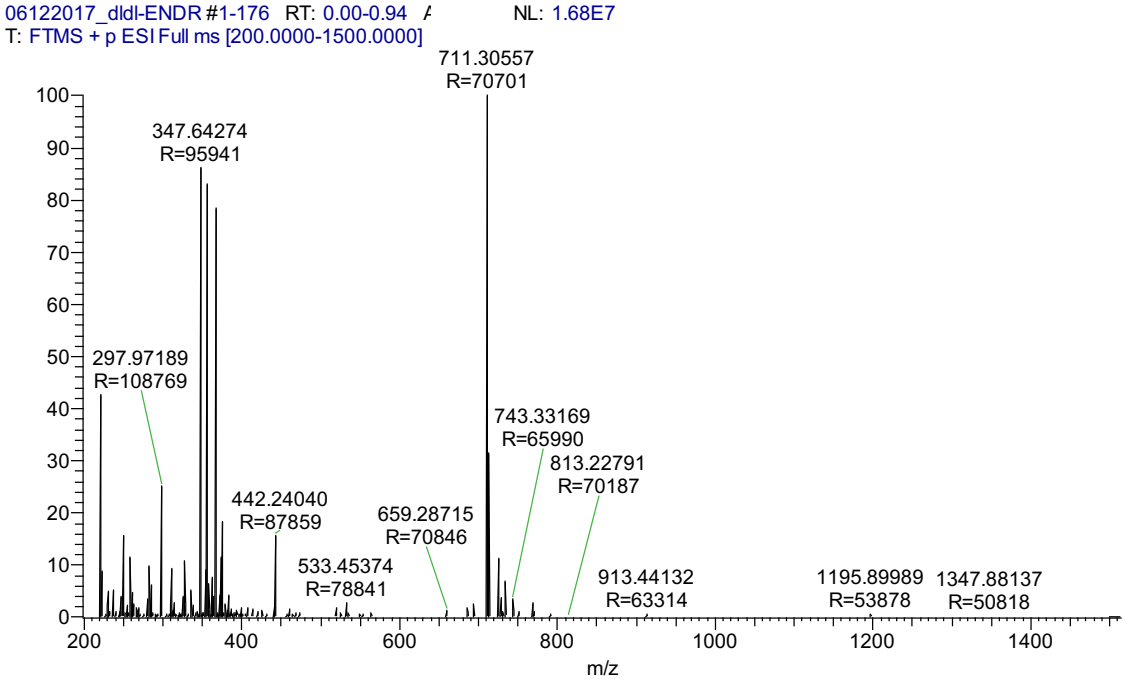
1H), 5.08 (d, *J* = 7.0 Hz, 1H), 4.66 (m, 1H), 4.35 (m, 1H), 4.29 (m, 1H), 4.22 (m, 1H), 4.09 (m, 1H), 3.67 (m, 1H), 3.31 (m, 4H), 3.11 (m, 1H), 2.86 (dd, *J* = 15.5, 7.1 Hz, 1H), 2.65 (m, 1H), 2.61-2.52 (m, 2H), 2.46-2.36 (m, 6H), 2.33 (m, 2H), 2.15 (m, 1H), 1.94 (m, 1H), 1.73 (m, 2H), 1.61 (m, 1H), 1.50 (m, 2H)



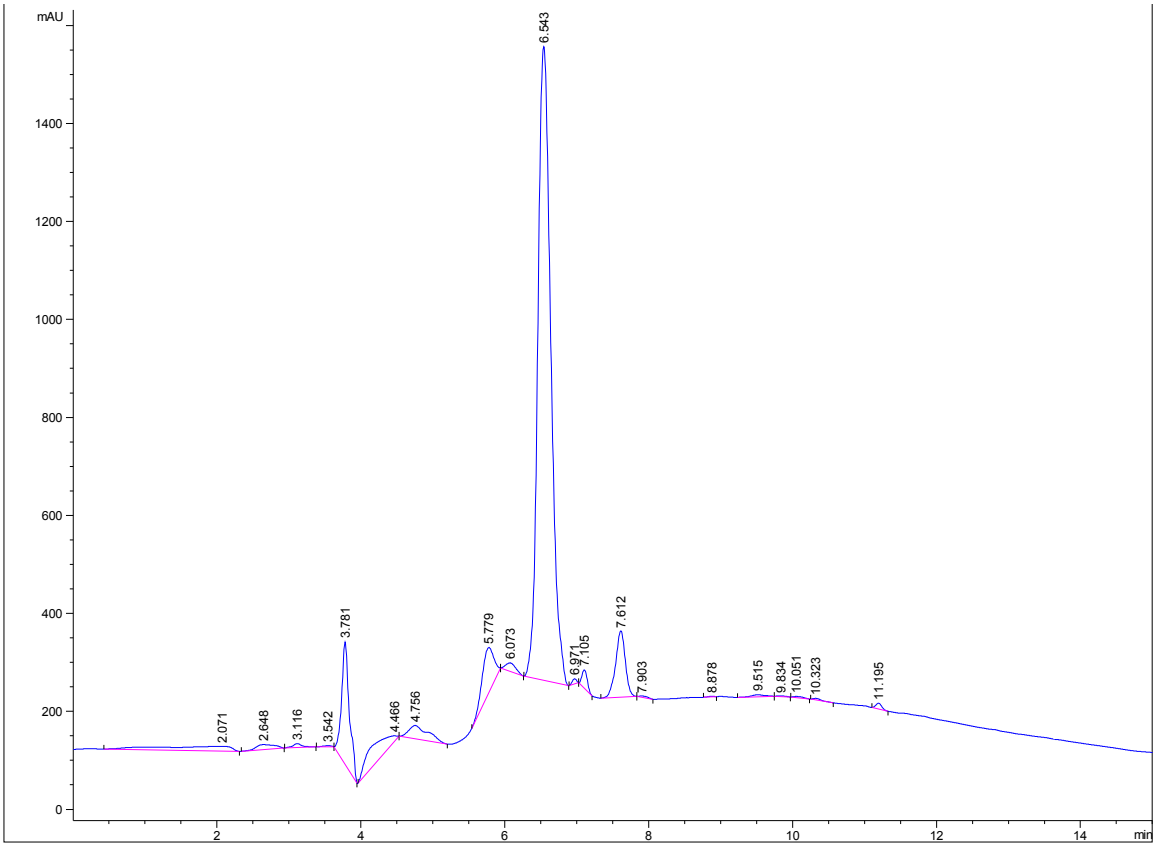
^{13}C NMR (100 MHz, DMSO-d_6) δ 175.6, 174.1, 174.0, 173.3, 171.9, 171.6, 170.4, 166.2, 156.7, 155.7, 65.3, 63.3, 55.1, 51.7, 48.5, 47.8, 47.2, 44.9, 42.3, 40.1, 33.9, 30.9, 29.1, 29.0, 28.1, 26.9, 24.7



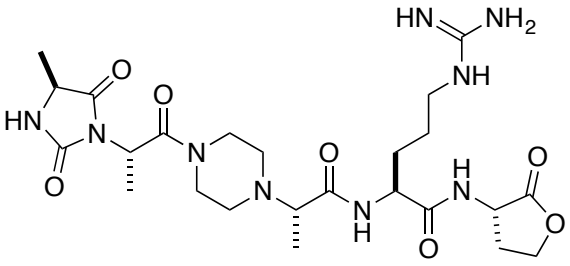
HRMS (ESI) m/z calcd for $C_{28}H_{43}N_{10}O_{12}^{+}$ 711.3056; found 711.3056 $(M+H)^{+}$



LC

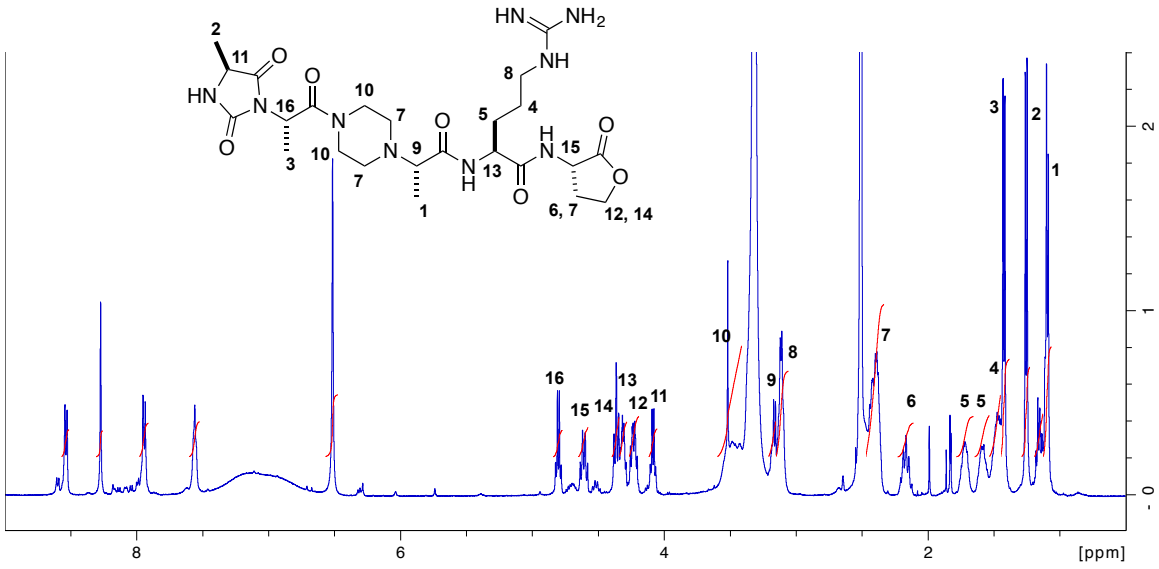


(*S*)-5-Guanidino-2-((*S*)-2-(4-((*S*)-2-((*S*)-4-methyl-2,5-dioxoimidazolidin-1-yl)propanoyl)piperazin-1-yl)propanamido)-*N*-((*S*)-2-oxotetrahydrofuran-3-yl)pentanamide (LLLL-**16aaar**)

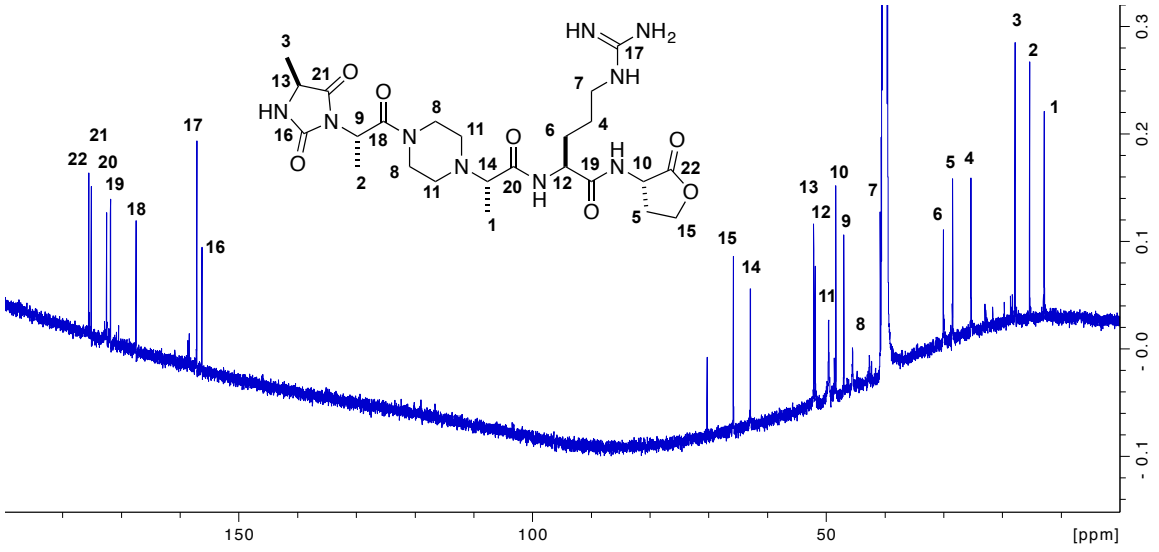


Data for LLLL-16aaar. ^{1}H NMR (400 MHz, DMSO-d_6) δ 8.54 (d, *J* = 8.0 Hz, 1H), 8.28 (s, 1H), 7.94 (d, *J* = 8.4 Hz, 1H), 7.56 (br, 1H), 6.52 (s, 2H), 4.80 (q, *J* = 7.1 Hz, 1H), 4.61 (q, *J* = 9.3 Hz, 1H), 4.36 (t, *J* = 8.7 Hz, 1H), 4.31 (m, 1H), 4.23 (m, 1H), 4.09 (q, *J* = 7.0

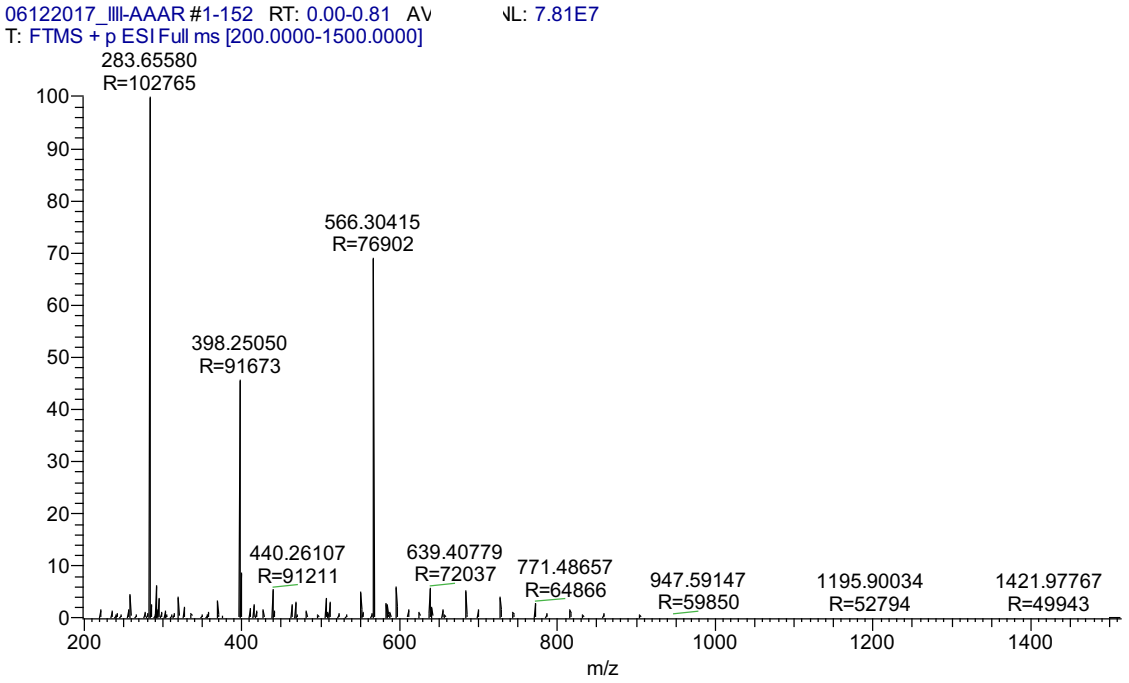
Hz, 1H), 3.60-3.41 (m, 4H), 3.17 (m, 1H), 3.12 (m, 2H), 2.41 (m, 5H), 2.17 (m, 1H), 1.72 (m, 1H), 1.59 (m, 1H), 1.48 (m, 2H), 1.42 (d, *J* = 7.1 Hz, 3H), 1.26 (d, *J* = 6.9 Hz, 3H), 1.09 (d, *J* = 6.9 Hz, 3H)



^{13}C NMR (100 MHz, DMSO-d_6) δ 175.0, 174.6, 172.0, 171.3, 167.0, 156.6, 155.7, 65.3, 62.4, 51.6, 51.3, 49.1, 47.8, 46.5, 45.0, 42.1, 40.3, 29.5, 27.9, 24.8, 17.3, 14.8, 12.3



HRMS (ESI) m/z calcd for $C_{24}H_{40}N_9O_7^+$ 566.3045; found 566.3042 $(M+H)^+$

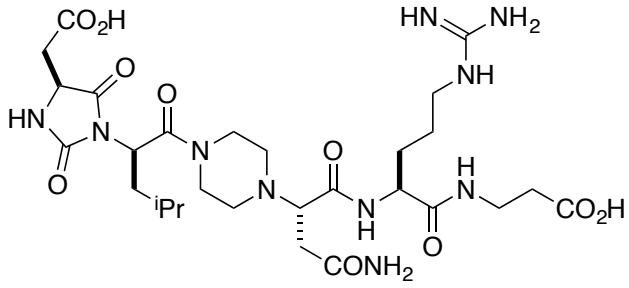


*Solid-phase Syntheses and Characterizations of Compounds **17** and **18***

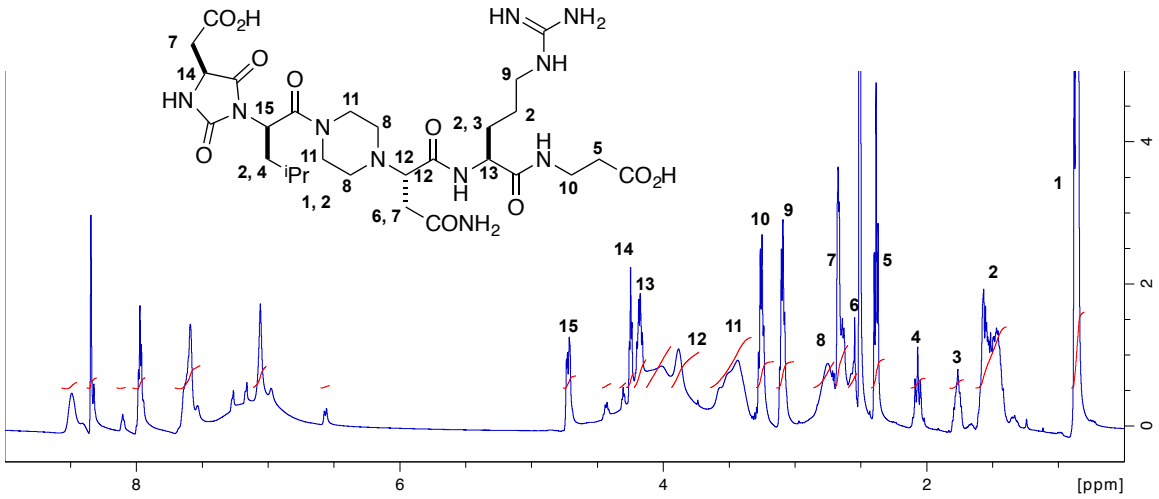
The derivatives of LDLL-**16dlnr**, compounds **17** and **18**, were synthesized by solid-phase and solution-phase syntheses. To synthesize the core structure, solid-phase synthesis was performed using 2-Cl-Trt resins loaded with Fmoc-β-Ala-OH. The protocols were the same as the syntheses of compounds **12** in Appendix C to obtain the protected-form of compound **17**. The crude product after cleaving with 20% HFIP/CH_2Cl_2 was reacted with the photoaffinity fragment **19** (1.1 eq), HBTU (1.1 eq) and $^{i}Pr_2NEt$ (3 eq) in DMF. The reaction was stirred at room temperature for 12 h. Solvent was completely removed, redissolved in CH_2Cl_2 and extracted with 10% citric acid, saturated $NaHCO_3$ and brine. The organic solution was collected, dried and removed in vacuo. Crude material was subsequently treated with TFA/Et_3SiH/water (95:2.5:2.5 v/v) for 3 h, solvent was removed with N_2 stream, and the crude material was purified by preparative

reverse-phase HPLC (10% - 50% MeCN/H_2O containing 0.05% TFA) to obtain compound **18** as white solids. For compound **17**, crude material after cleavage from beads was treated with TFA/Et_3SiH/water cocktail and purified by preparative reverse-phase HPLC (10% - 50% MeCN/water containing 0.05% TFA).

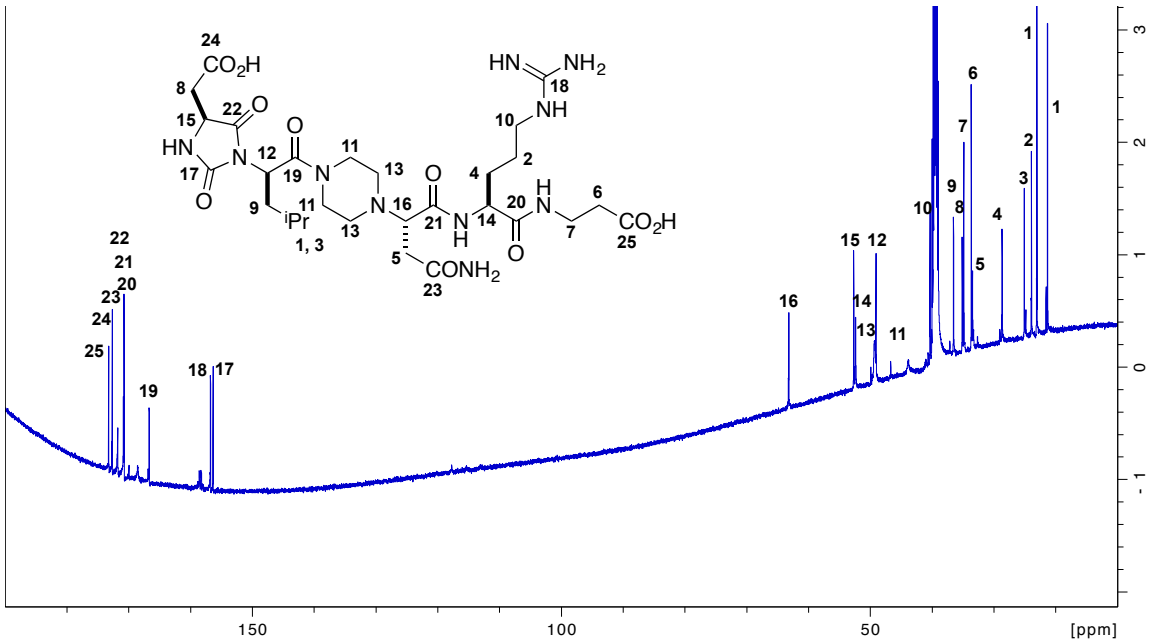
3-((*S*)-2-((*S*)-4-Amino-2-(4-((*R*)-2-((*S*)-4-(carboxymethyl)-2,5-dioxoimidazolidin-1-yl)-4-methylpentanoyl)piperazin-1-yl)-4-oxobutanamido)-5-guanidinopentanamido)propanoic acid (**17**)



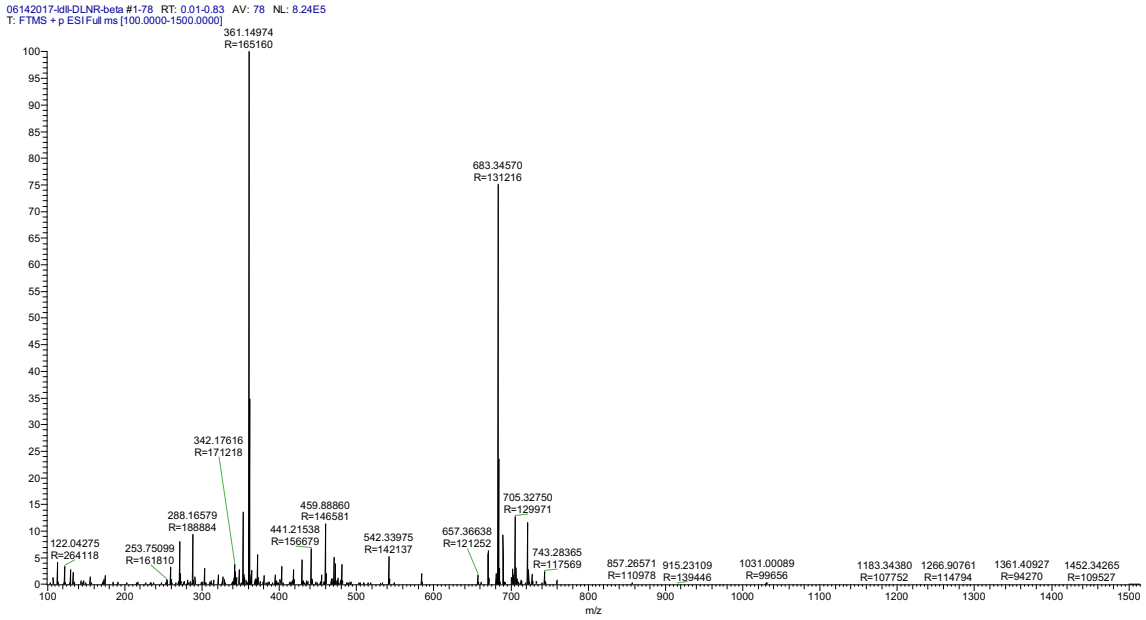
Data for Compound 17. ^{1}H NMR (400 MHz, DMSO-d_6) δ 8.49 (br, 1H), 8.34 (m, 1H), 7.98 (m, 1H), 7.62 (br, 2H), 7.06 (br, 2H), 4.72 (m, 1H), 4.25 (m, 1H), 4.18 (m, 1H), 3.89 (m, 1H), 3.65-3.33 (m, 4H), 3.26 (m, 2H), 3.10 (m, 2H), 2.86-2.70 (m, 4H), 2.70-2.60 (m, 3H), 2.60-2.53 (m, 1H), 2.38 (t, *J* = 7.2 Hz, 2H), 2.07 (m, 1H), 1.76 (m, 1H), 1.62-1.40 (m, 5H), 0.87 (m, 6H);



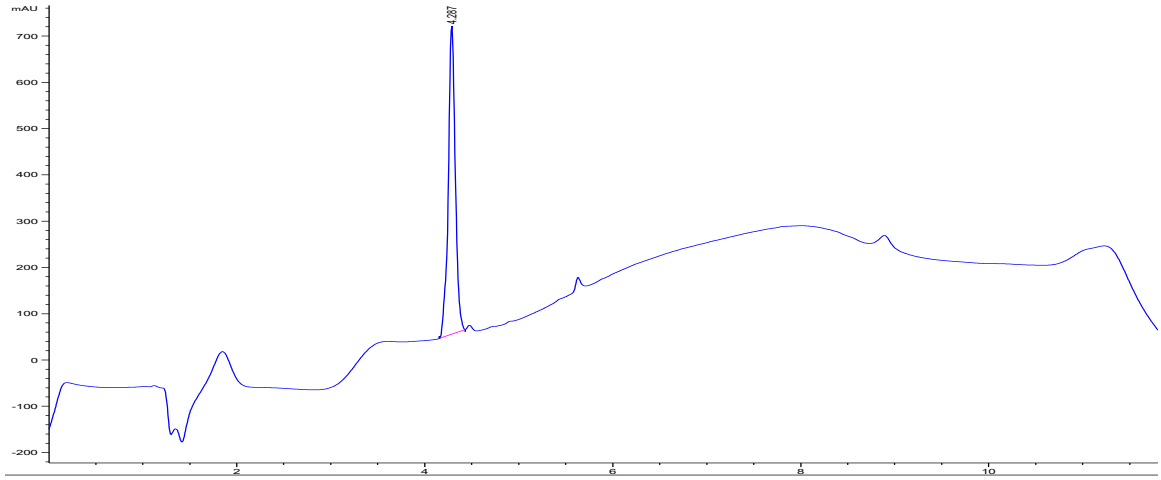
^{13}C NMR (100 MHz, DMSO-d_6) δ 173.2, 172.8, 172.7, 171.8, 170.9, 170.8, 166.7, 156.8, 156.4, 63.2, 52.7, 52.4, 49.4, 49.1, 43.9, 41.1, 40.4, 36.6, 35.2, 34.9, 33.7, 33.5, 28.7, 25.1, 24.0, 23.1, 21.3



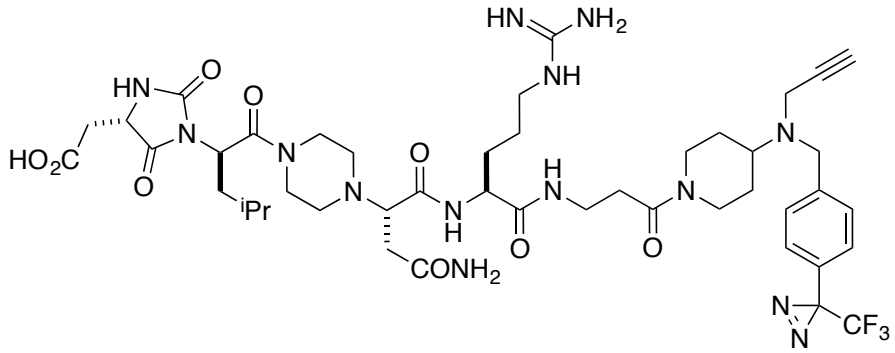
HRMS (ESI) m/z calcd for $C_{28}H_{47}N_{10}O_{10}^{+}$ 683.3471; found 683.3457 $(M+H)^{+}$



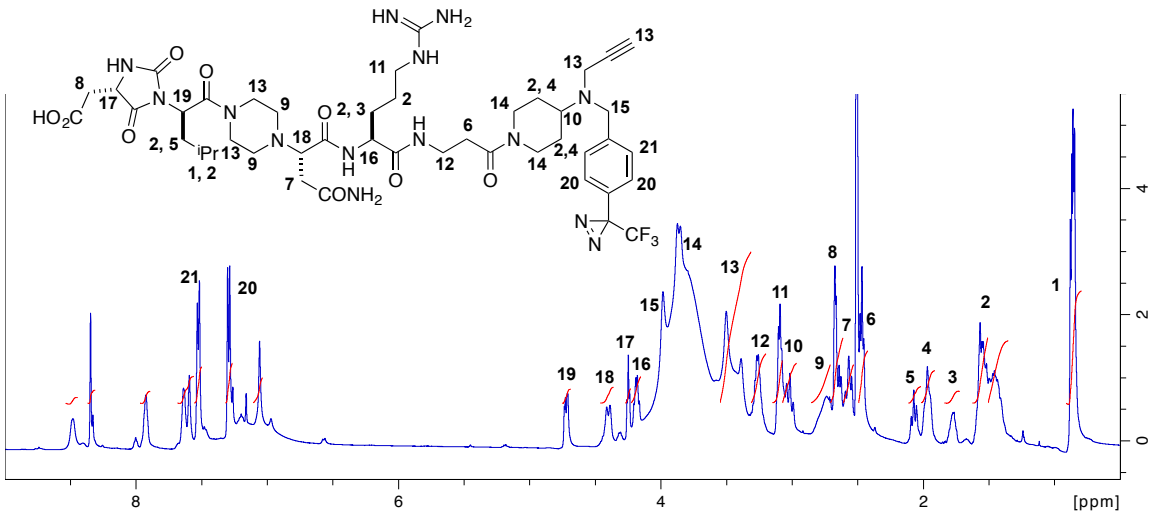
LC



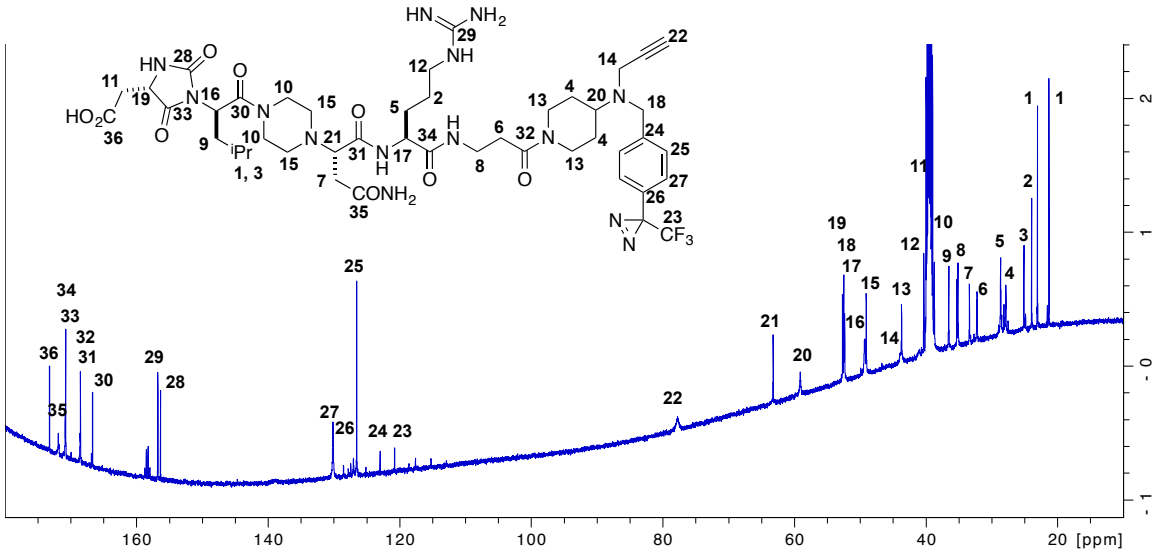
2-((*S*)-1-((*R*)-1-(4-((*S*)-4-Amino-1-(((*S*)-5-guanidino-1-oxo-1-((3-oxo-3-(4-(prop-2-yn-1-yl(4-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)phenyl)amino)piperidin-1-yl)propyl)amino)pentan-2-yl)amino)-1,4-dioxobutan-2-yl)piperazin-1-yl)-4-methyl-1-oxopentan-2-yl)-2,5-dioxoimidazolidin-4-yl)acetic acid (**18**)



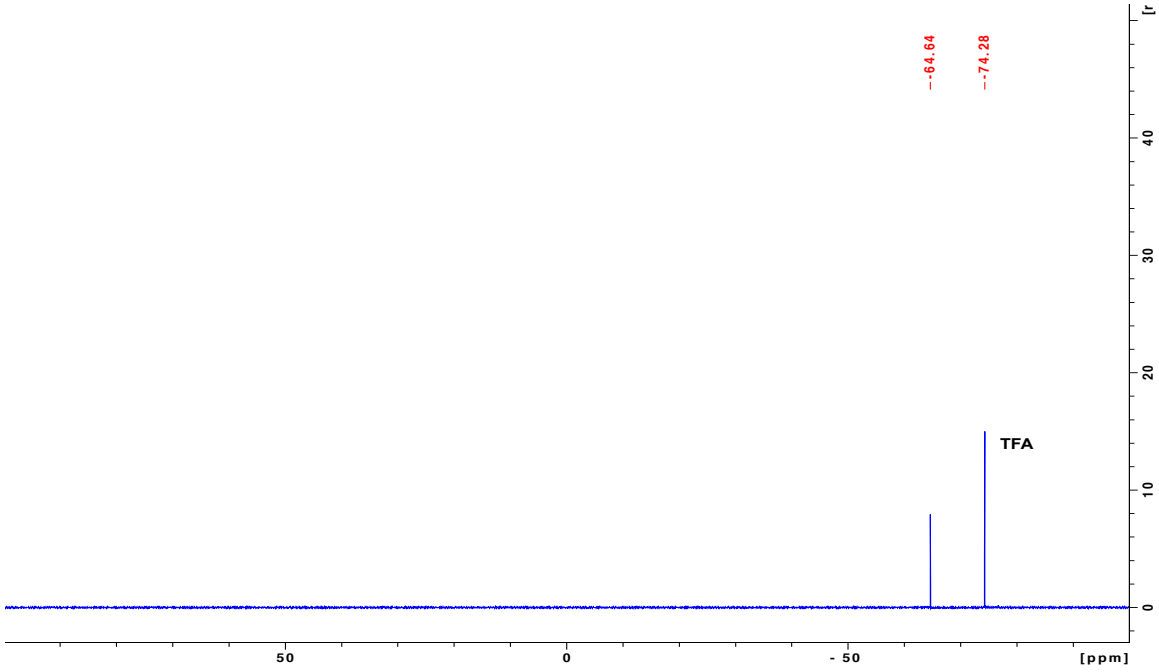
Data for Compound 18. ^{1}H NMR (400 MHz, DMSO-d$_6$) δ 8.49 (br, 1H), 8.34 (m, 1H), 7.93 (br, 1H), 7.66-7.57 (br, 2H), 7.53 (d, *J* = 7.7 Hz, 2H), 7.30 (d, *J* = 7.9 Hz, 2H), 7.06 (br, 2H), 4.72 (m, 1H), 4.40 (m, 1H), 4.25 (m, 1H), 4.19 (m, 1H), 3.99 (s, 2H), 3.85 (m, 4H), 3.55-3.32 (m, 7H), 3.26 (m, 2H), 3.10 (m, 2H), 3.02 (m, 1H), 2.85-2.70 (m, 4H), 2.68 (m, 2H), 2.63 (m, 2H), 2.47 (m, 2H), 2.07 (m, 1H), 1.97 (m, 2H), 1.78 (m, 1H), 1.61-1.38 (m, 7H), 0.86 (m, 6H)



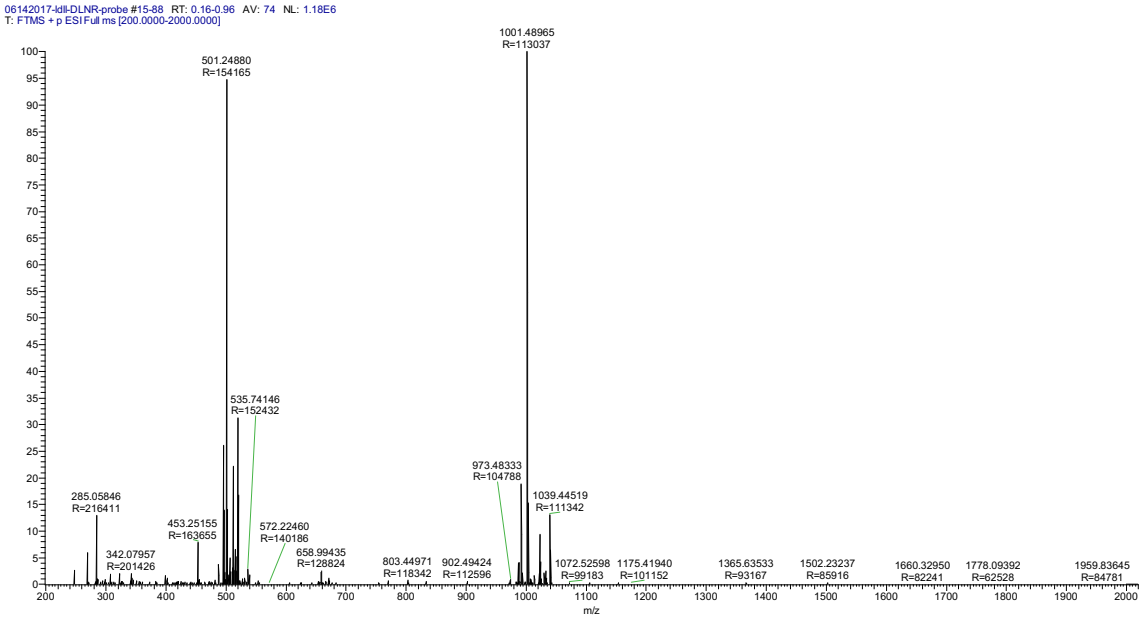
^{13}C NMR (100 MHz, DMSO-d_6) δ 173.3, 171.9, 170.9, 170.8, 168.6, 168.5, 166.7, 156.8, 156.4, 130.1, 127.0, 126.6, 123.0, 120.9, 77.8, 63.3, 59.1, 52.7, 52.5, 52.4, 49.4, 49.1, 43.9, 43.7, 41.1, 40.3, 38.7, 36.6, 35.4, 35.1, 33.4, 32.3, 28.7, 27.9, 25.1, 24.0, 23.1, 21.3



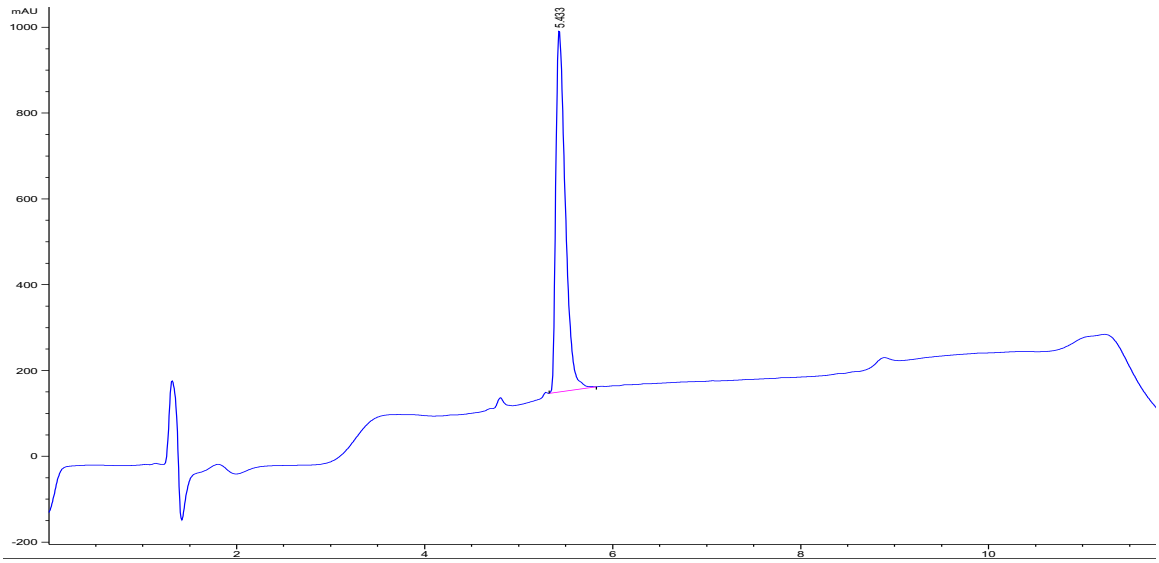
^{19}F NMR spectrum



HRMS (ESI) m/z calcd for $C_{45}H_{64}F_3N_{14}O_9^+$ 1001.4927; found 1001.4897 $(M+H)^+$



LC



*Solution-phase Syntheses of LDLL-**16dlnr***

Synthesis of L-homoserine lactone hydrobromide salt

L-methionine (2.98 g, 20 mmol) and bromoacetic acid (3.06 g, 22 mmol) were dissolved in ^{i}PrOH, H_2O and AcOH (5:5:2, 30 mL). Mixture was stirred in the reflux condition for 12 h, then solvent was completely removed under vacuum to obtain dark brown oil. This crude was dissolved in 4:1 ^{i}PrOH:HBr in 33 % AcOH, and let it cool in the fridge to induce white precipitate. Solid was filtered and washed with ^{i}PrOH. The supernatant was concentrated in vacuo and reprecipitated again with the same condition. Precipitation was repeated until no further solid formed. All crops of precipitate were combined to obtain the product as a hydrobromide salt (2.76 g, 75%). NMR spectra were identical to the reference.[179]

*Synthesis of dipeptide LD-**7d'l***

L-Fmoc-Asp(^{t}Bu)-OH (1.94 g, 4.72 mmol), EDCI (1.00 g, 5.19 mmol) and HOBt (0.70 g, 5.19 mmol) were stirred in 50 mL CH_2Cl_2 at 0 ºC under N_2 atmosphere for 30 min. $^{i}Pr_2NEt$ (2.5 mL, 14.2 mmol) was added followed by D-Leu-OBn (1.04 g, 4.72 mmol). The reaction was allowed to stir at room temperature for 12 h, then was diluted with 100 mL CH_2Cl_2 and extracted three times with 10% citric acid, three times with saturated $NaHCO_3$ and brine. Organic phase was dried with $MgSO_4$ and evaporated in vacuo to obtain the product as a white solid (2.67 g, 92%), which was used in the next step without further purification.

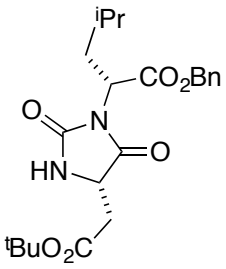
*Synthesis of hydantoin LD-**8d'l***

Dipeptide LD-**7d'l** (0.62 g, 1.00 mmol) was stirred in Et_2NH/CH_2Cl_2 solution (4 mL, 1:1 v/v) for 3 h, solvent was removed in vacuo, and then the crude material was purified by flash column chromatography using 2-5% $MeOH/CH_2Cl_2$ with 0.1% Et_3N to obtain product as a clear oil. Purified NH_2-dipeptide (0.40 g, 1.00 mmol) and $NaHCO_3$ (0.25 g, 3.03 mmol) were dissolved in 20 mL MeCN. *p*-Nitrophenyl chloroformate (0.21 g, 1.1 mmol) was added and the reaction was stirred at room temperature for 3 h under N_2 atmosphere. Water (~ 10 mL) was added to produce bright yellow solution, which was stirred for additional 12 h at room temperature. Organic solvent was removed in vacuo, then EtOAc (100 mL) was added into the aqueous solution. The organic layer was separated, and extracted with 5% Na_2CO_3 five times to remove *p*-nitrophenol byproduct. The clear organic solution was further dried with brine and $MgSO_4$. The mixture was reduced in vacuo to obtain crude oil, which was purified by flash column chromatography using EtOAc:hexanes (1:4 v/v) to obtain product as a white solid (0.34 g, 80%)

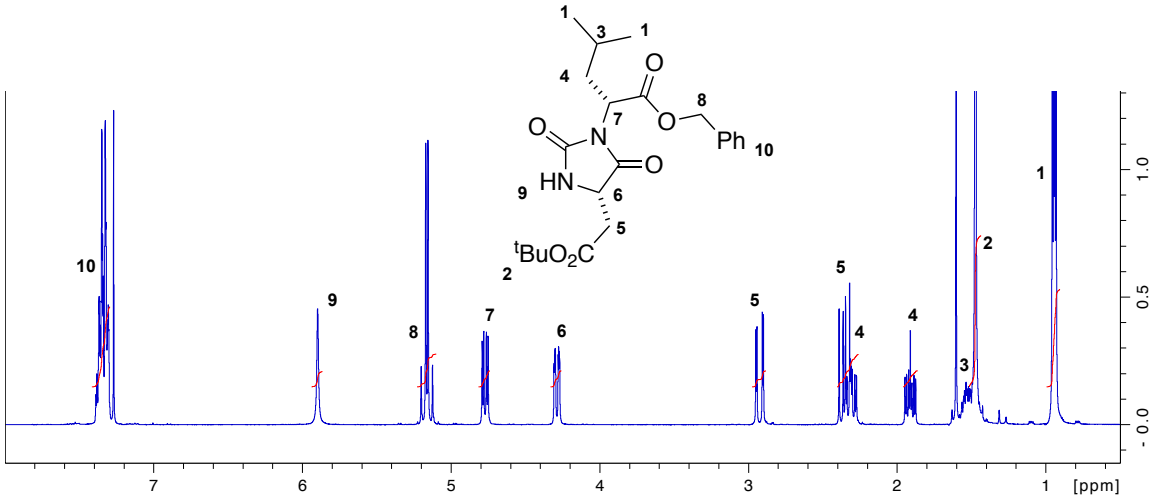
*Benzyl ester deprotection (LD-**20d'l**)*

Hydantoin LD-**8d'l** (2.51 g, 6 mmol) and 10% Pd/C (620 mg, 0.6 mmol) were purged under N_2 before adding 50 mL MeOH. H_2 balloon was purged for 10 min, and the reaction was stirred under H_2 atmosphere for 12 h. The mixture was filtered through celite and the clear solution was reduced in vacuo to obtain compound LD-**20d'l** (1.99 g, 100%) as clear oil.

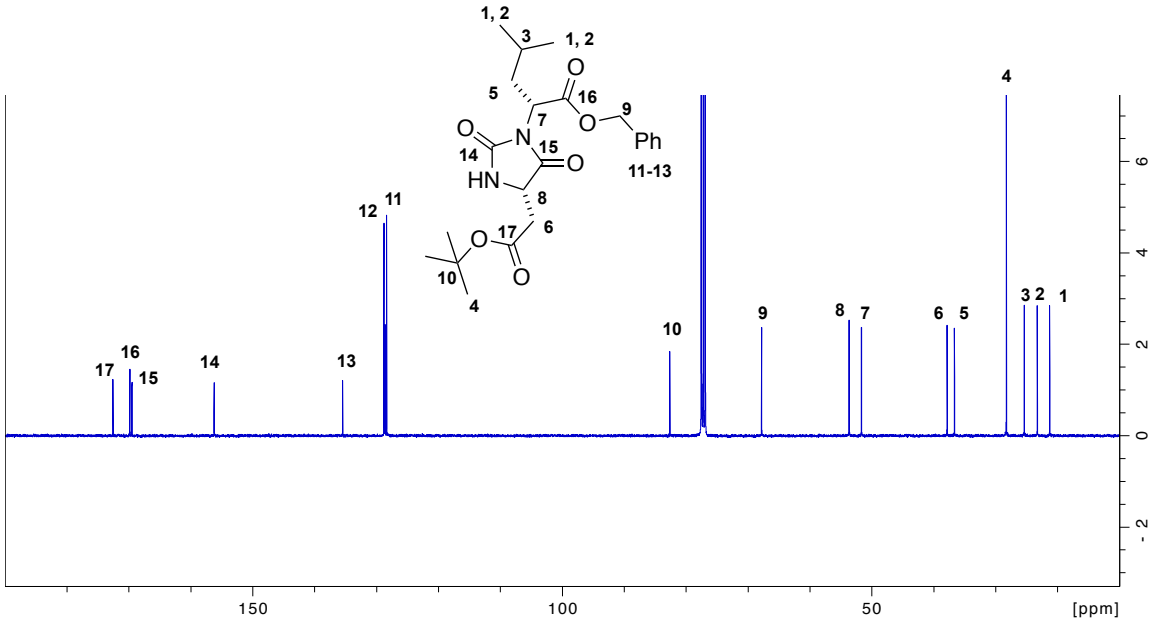
Benzyl (*R*)-2-((*S*)-4-(2-(*tert*-butoxy)-2-oxoethyl)-2,5-dioxoimidazolidin-1-yl)-4-methylpentanoate (LD-**8d'l**)



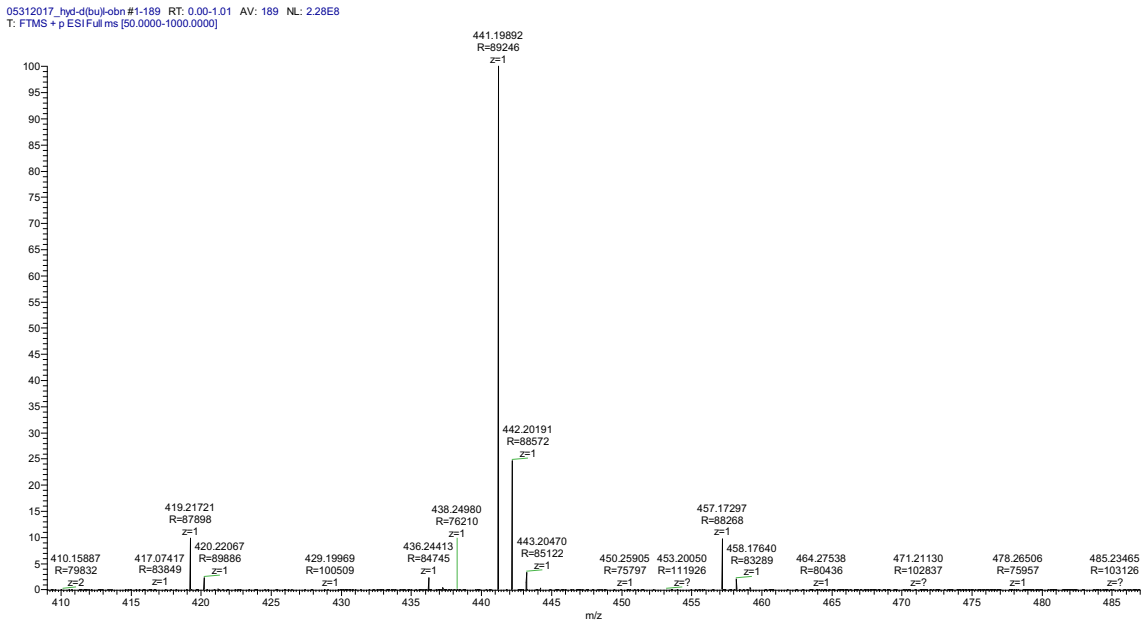
Data for LD-8d'l. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.39-7.30 (m, 5H), 5.90 (br, 1H), 5.17 (d, *J* = 6.0 Hz, 2H), 4.77 (dd, *J* = 11.5, 4.4 Hz, 1H) 4.29 (m, 1H), 2.93 (dd, *J* = 17.4, 2.8 Hz, 1H), 2.39-2.27 (m, 2H), 1.96-1.87 (m, 1H), 1.58-1.50 (m, 1H), 1.48 (s, 9H) 0.94 (dd, *J* = 6.6, 3.6 Hz, 6H)



^{13}C NMR (100 MHz, $CDCl_3$) δ 172.6, 169.8, 169.5, 156.2, 135.5, 128.8, 128.6, 128.4, 82.6, 67.8, 53.7, 51.7, 37.8, 36.7, 28.3, 25.4, 23.3, 21.3



HRMS (ESI) m/z calcd for $C_{22}H_{30}N_2NaO_6Na^+$ 441.1996; found 441.1989 $(M+Na)^+$



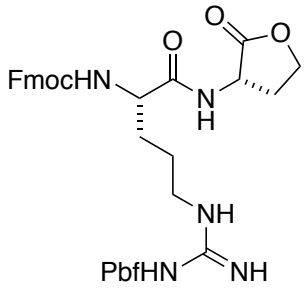
General Procedure of Amide Coupling with HBTU

N-protected amino acid (1 eq) was activated with HBTU (1.1 eq) and $^{i}Pr_2NEt$ (1 eq) in DMF (0.2 M) at 0 ºC for 10 min. Then, amine compound (1 eq) was added followed by $^{i}Pr_2NEt$ (2.5 eq). Reaction was stirred at ambient temperature for 12 h under N_2 atmosphere. DMF was removed in vacuo, then crude material was redissolved in EtOAc and extracted with 10% citric acid, saturated $NaHCO_3$ and brine. After drying with $MgSO_4$, the compound was purified by flash column chromatography using 5% MeOH/CH_2Cl_2 to obtain product.

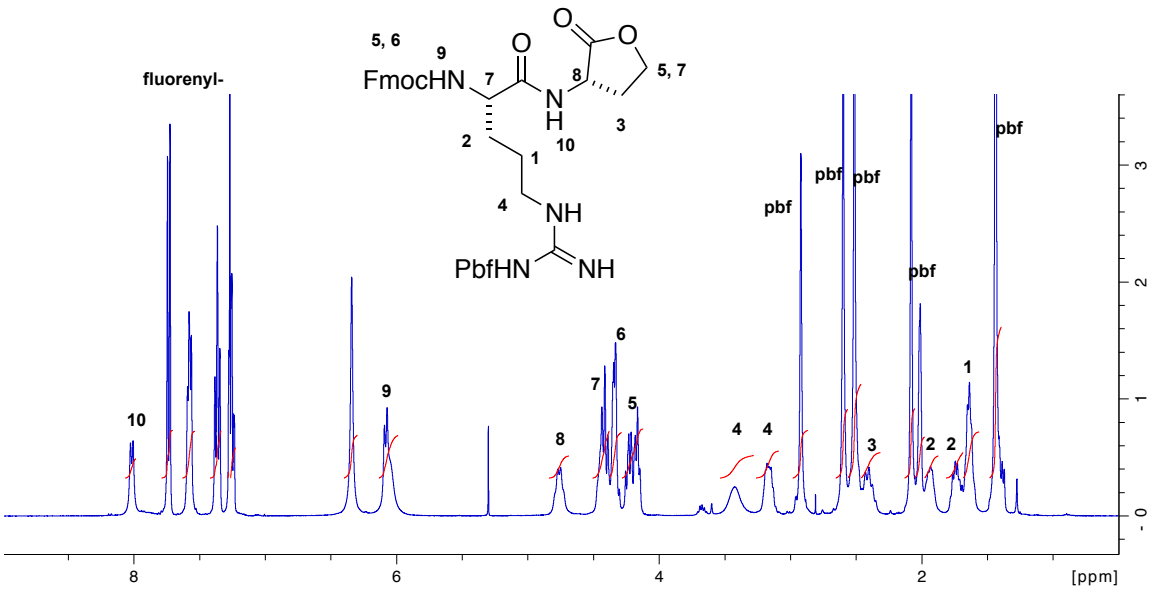
Nosyl deproection

Compound LL-**22n'r'** (0.80 g, 0.71 mmol) was dissolved in 7 mL DMF, then mercaptoethanol (0.10 mL, 1.43 mmol) and DBU (0.21 mmol, 1.43 mmol) were added subsequently, and the reaction was stirred for 2 h. After the reaction completed, the mixture was diluted with EtOAc and extracted 3 times with saturated $NaHCO_3$. The organic phase was dried over Na_2SO_4 and removed in vacuo to obtain crude oil. Small amount of CH_2Cl_2 was added to dissolve oil, then cold *tert*-butyl methyl ether was added to induce precipitation. The solid was filtered, washed with ether to obtain product LL-**23n'r'** as a solid (0.64 g, 96%).

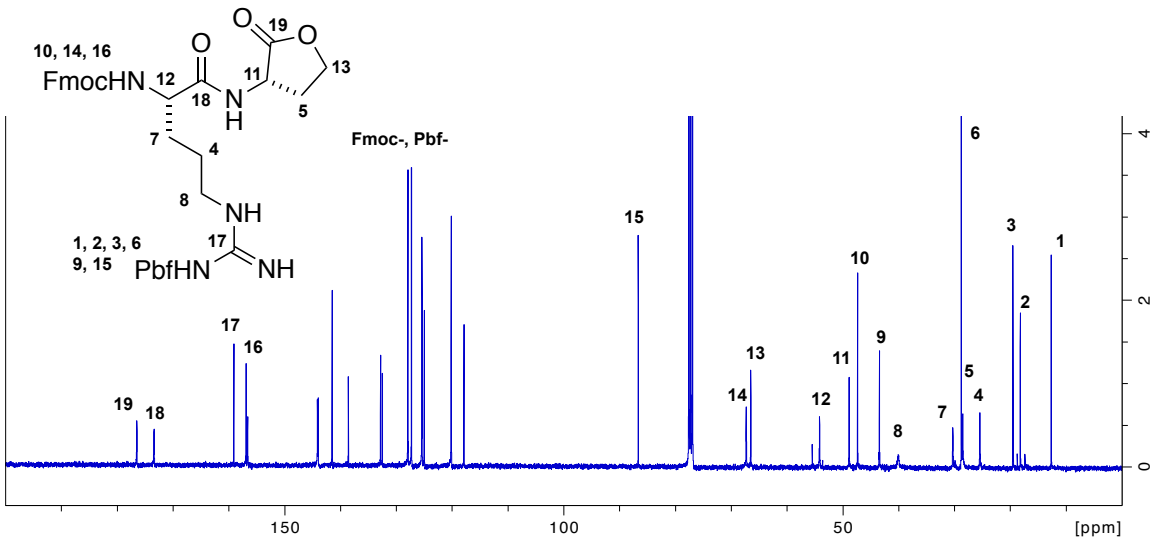
(9*H*-Fluoren-9-yl)methyl ((*S*)-1-oxo-1-(((*S*)-2-oxotetrahydrofuran-3-yl)amino)-5-(3-((2,2,4,6,7-pentamethyl-2,3-dihydrobenzofuran-5-yl)sulfonyl)guanidino)pentan-2-yl)carbamate (L-**21r'**)



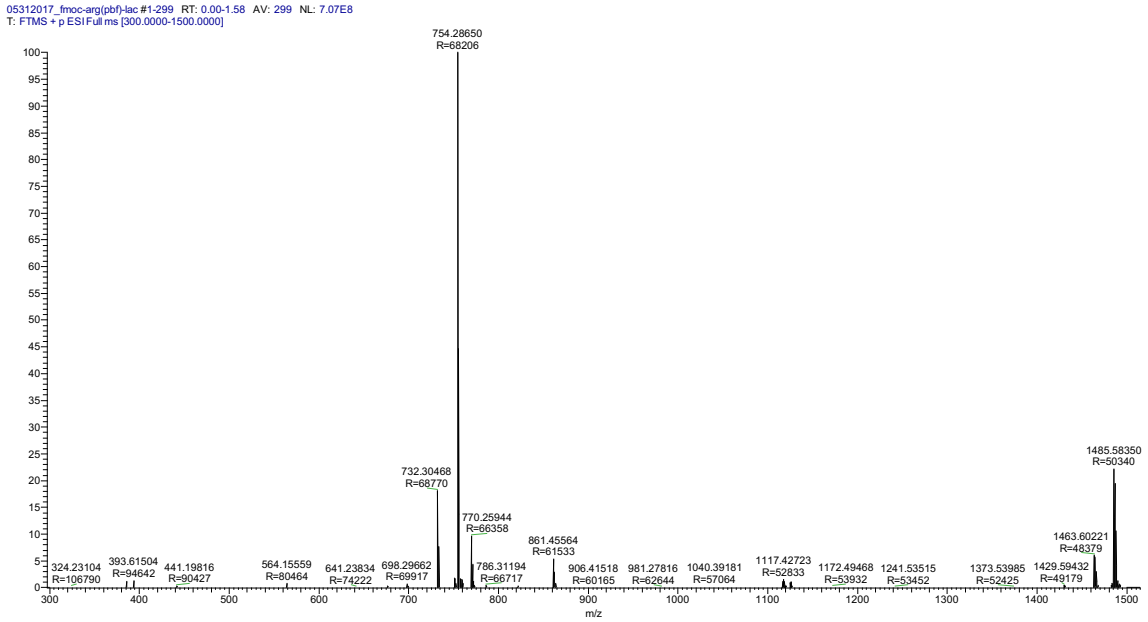
Data for L-21r'. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.02 (br, 1H), 7.74 (d, *J* = 7.7 Hz, 2H), 7.58 (m, 2H), 7.36 (t, *J* = 7.4 Hz, 2H), 7.30-7.22 (m, 2H), 6.34 (br, 1H), 6.07 (br, 2H), 4.76 (m, 1H), 4.49-4.37 (m, 2H), 4.37-4.29 (m, 2H), 4.27-4.12 (m, 2H), 3.43 (m, 1H), 3.17 (m, 1H), 2.92 (s, 2H), 2.60 (s, 3H), 2.51 (m, 4H), 2.46-2.31 (m, 1H), 2.08 (s, 3H), 1.98-1.88 (m, 1H), 1.80-1.68 (m, 1H), 1.68-1.58 (m, 2H), 1.44 (m, 7H)



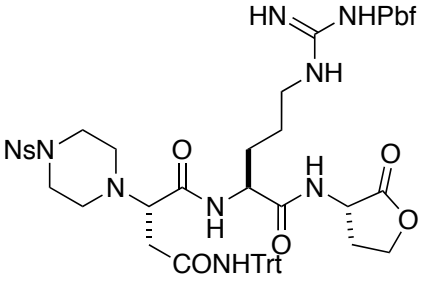
^{13}C NMR (100 MHz, $CDCl_3$) δ 176.5, 173.4, 159.1, 156.9, 156.6, 144.1, 144.0, 141.5, 138.6, 132.8, 132.5, 127.9, 127.3, 125.4, 124.9, 120.1, 117.8, 86.6, 67.3, 66.5, 54.2, 48.9, 47.3, 43.4, 40.1, 30.3, 28.8, 28.5, 25.4, 19.5, 18.2, 12.7, 12.6



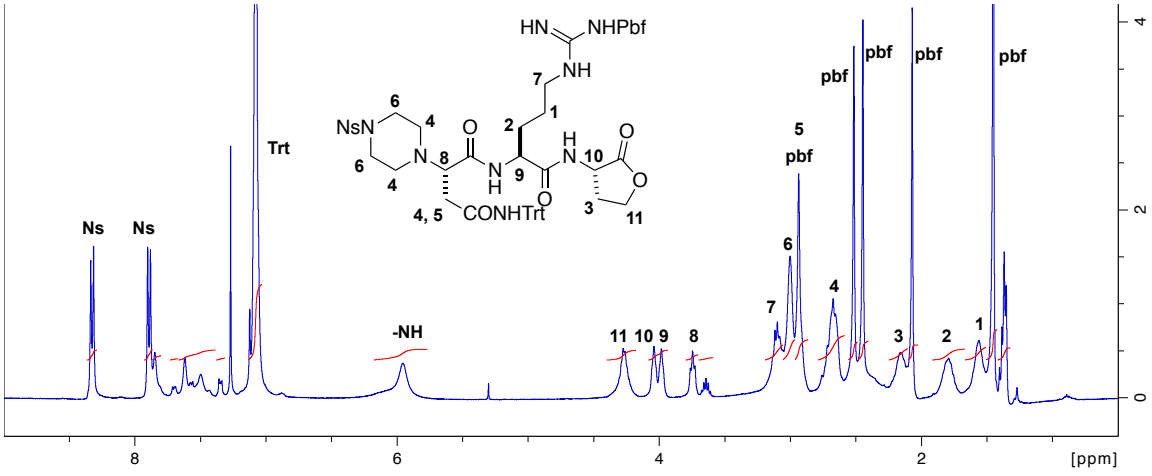
HRMS (ESI) m/z calcd for $C_{38}H_{46}N_5O_8S^+$ 732.3062; found 732.3047 $(M+H)^+$



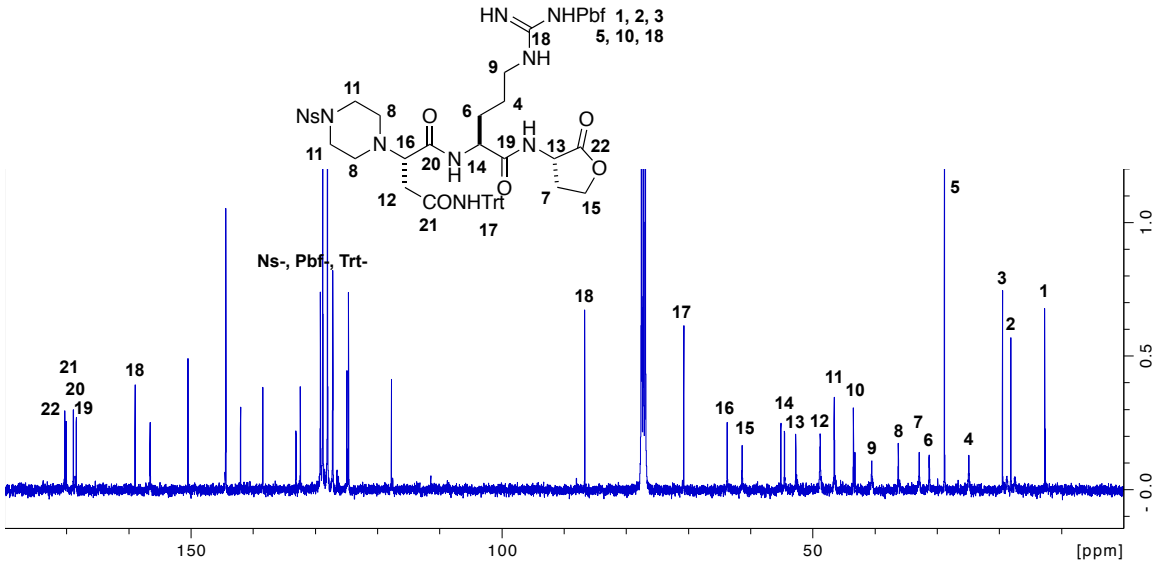
(*S*)-2-(4-((4-Nitrophenyl)sulfonyl)piperazin-1-yl)-*N*^{1}-((*S*)-1-oxo-1-(((*S*)-2-oxotetrahydrofuran-3-yl)amino)-5-(3-((2,2,4,6,7-pentamethyl-2,3-dihydrobenzofuran-5-yl)sulfonyl)guanidino)pentan-2-yl)-*N*^{4}-tritylsuccinamide (LL-**22n'r'**)



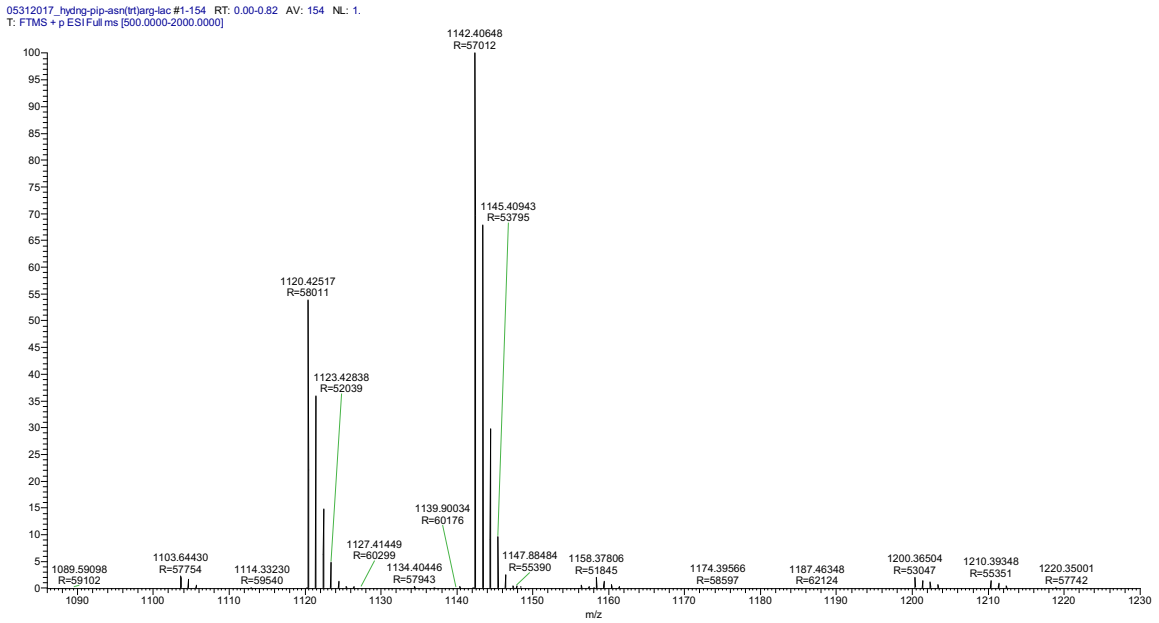
Data for LL-22n'r'. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.33 (d, *J* = 8.7 Hz, 2H), 7.89 (d, *J* = 8.6 Hz, 2H), 7.08 (s, 15H), 5.96 (br, 2H), 4.27 (m, 2H), 4.01 (m, 2H), 3.74 (m, 1H), 3.16-3.04 (m, 3H), 3.00 (m, 4H), 2.94 (m, 4H), 2.78-2.59 (m, 5H), 2.51 - 2.44 (m, 8H), 2.22-2.10 (m, 2H), 2.08 (s, 3H), 1.79 (m, 2H), 1.56 (m, 2H), 1.45 (s, 6H)



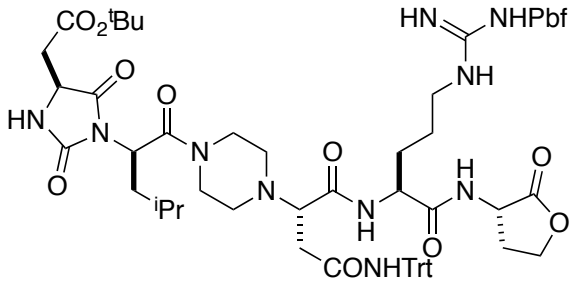
^{13}C NMR (100 MHz, $CDCl_3$) δ 170.3, 170.1, 168.9, 168.4, 159.0, 156.6, 150.5, 144.4, 142.0, 138.4, 133.1, 132.4, 129.2, 128.8, 128.0, 127.2, 124.9, 124.7, 117.7, 86.7, 70.7, 63.8, 61.3, 55.1, 54.5, 52.7, 48.8, 46.5, 43.5, 43.2, 40.5, 36.2, 32.9, 31.3, 28.8, 24.9, 19.5, 18.1, 12.7, 12.6



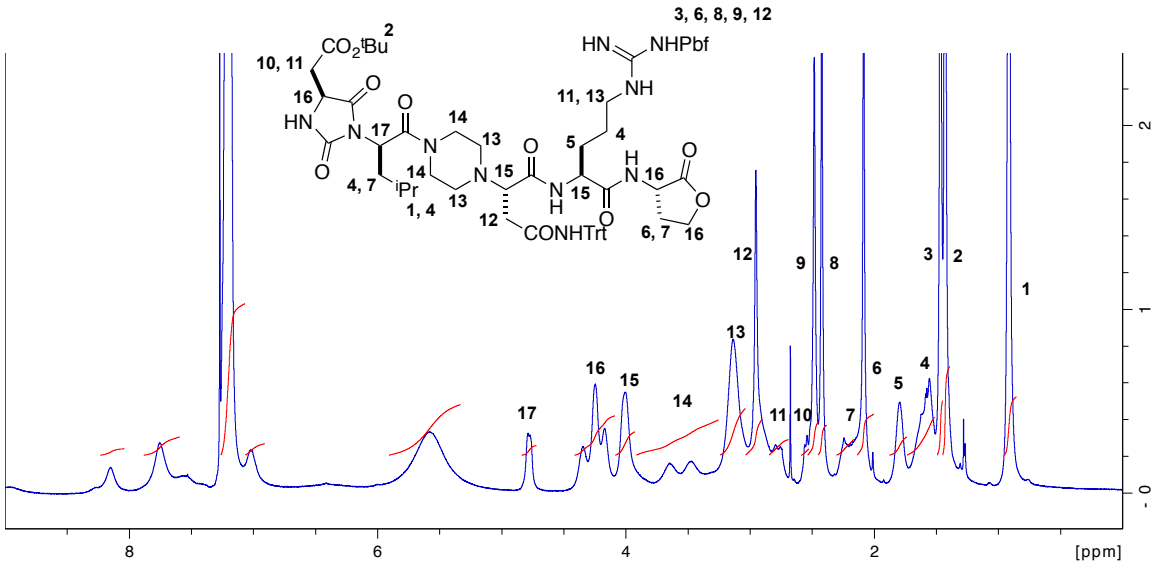
HRMS (ESI) m/z calcd for $C_{56}H_{66}N_9O_{12}S_2^+$ 1120.4267; found 1120.4252 $(M+H)^+$



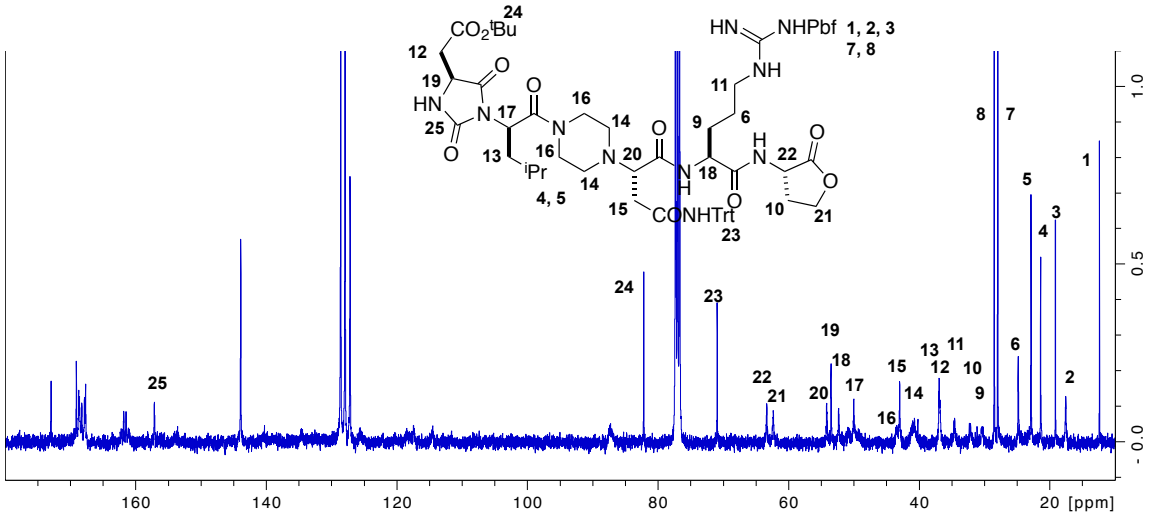
tert-Butyl 2-((*S*)-1-((*R*)-1-(4-((*S*)-1,4-dioxo-1-(((*S*)-1-oxo-1-(((*S*)-2-oxotetrahydrofuran-3-yl)amino)-5-(3-((2,2,4,6,7-pentamethyl-2,3-dihydrobenzofuran-5-yl)sulfonyl)guanidino)pentan-2-yl)amino)-4-(tritylamino)butan-2-yl)piperazin-1-yl)-4-methyl-1-oxopentan-2-yl)-2,5-dioxoimidazolidin-4-yl)acetate (LDLL-**16d'ln'r'**)



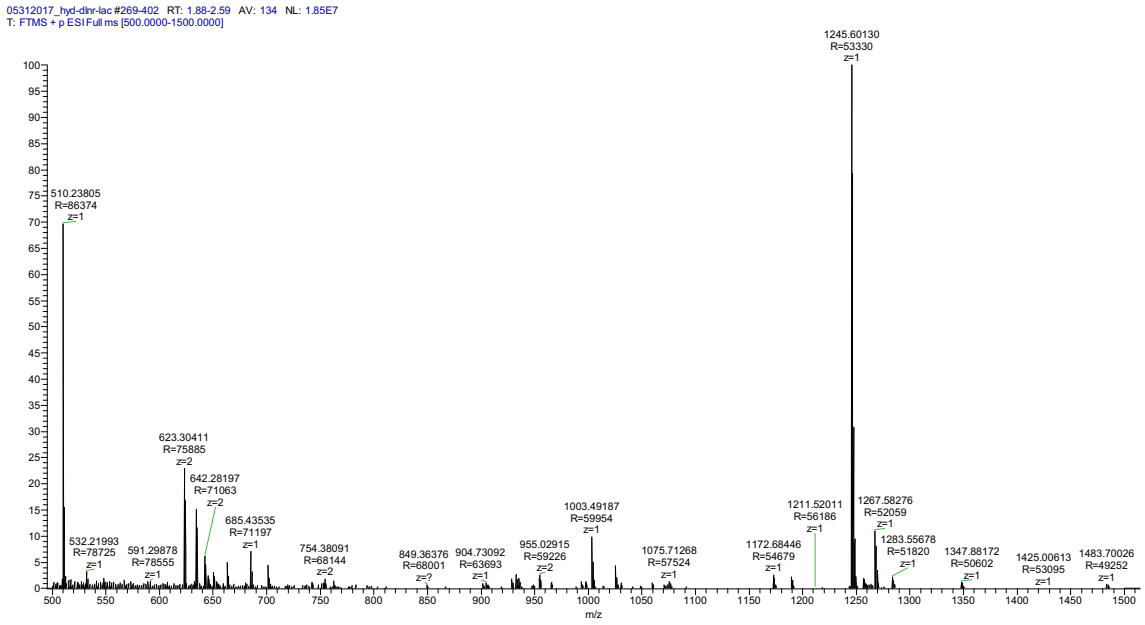
Data for LDLL-16d'ln'r'. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.15 (br, 1H), 7.75 (br, 2H), 7.20 (m, 15H), 7.02 (m, 1H), 5.58 (m, 3H), 4.79 (m, 1H), 4.26 (m, 4H), 4.01 (m, 2H), 3.79-3.39 (m, 4H), 3.14 (m, 5H), 2.96 (m, 4H), 2.84-2.69 (m, 2H), 2.55 (m, 1H), 2.50 (s, 3H), 2.43 (s, 3H), 2.21 (m, 2H), 2.09 (m, 4H), 1.80 (m, 2H), 1.71-1.51 (m, 4H), 1.47 (m, 6H), 1.43 (s, 9H) 0.92 (m, 6H);



^{13}C NMR (100 MHz, $CDCl_3$) δ 173.0, 169.1, 168.7, 168.6, 168.3, 167.8, 167.7, 161.8, 161.5, 157.1, 143.9, 128.6, 127.9, 127.2, 82.2, 70.9, 63.4, 62.4, 54.2, 53.5, 52.3, 50.0, 43.5, 43.3, 40.7, 37.0, 36.9, 34.6, 32.3, 30.4, 28.5, 28.0, 24.8, 22.9, 21.4, 19.1, 17.6, 12.4



HRMS (ESI) m/z calcd for $C_{65}H_{85}N_{10}O_{13}S^+$ 1245.6013; found 1245.6013 $(M+H)^+$



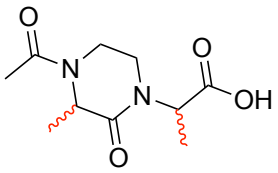
APPENDIX E

SUPPORTING INFORMATION FOR CHAPTER V

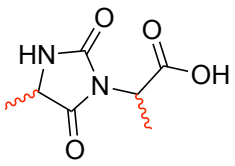
Matching Analyses of Dipeptide Fragments on Ideal Secondary Structures

All the mimics featured in Chapter V may be considered as modular, made up of dipeptide mimics connected to another fragment that projects a third amino acid. We decided to compare common dipeptide fragments overlaid on ideal secondary structures. To facilitate this, the algorithm at the heart of EKOS was modified to evaluate overlays on dipeptide mimics.

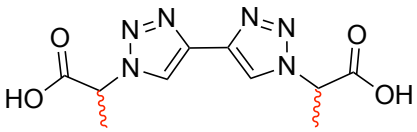
All the mimics overlaid on all the secondary structures with an RMSD of <1.2 Å, except **37** and **39** that did not match well with turns and β-strands. Consequently, RMSD scores for the best overlays for each mimic (over all isomers) were averaged to give the data in Figure E.1a, but matches for **37** and **39** on turns and β-strands were not considered. These data indicate **35** and **36** were the best dipeptide mimics in the series (Figure E.1). Mimic **42** is interesting because it is the core of the so-called "triazolamers", and the best dipeptide mimic for helical structures was **38**. Conversely, there could be clear reasons to anticipate other mimics in the series would *not* overlay well. For instance, **44** has an inappropriate *C*α - *C*α separation, while mimic **45** is locked in a *cis*-amide bond conformation. Consequently, henceforth the discussion of dipeptide mimics is restricted to **35**, **36**, **38**, and **42**.



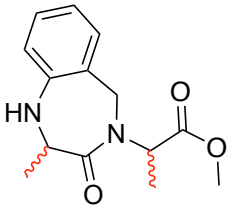
35



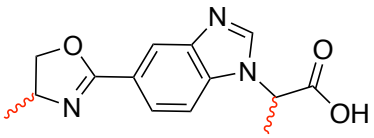
36



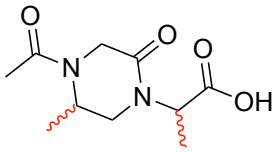
37



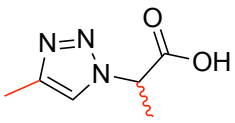
38



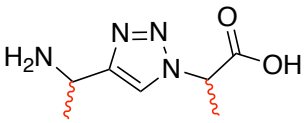
39



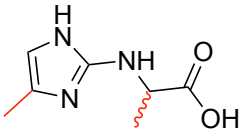
40



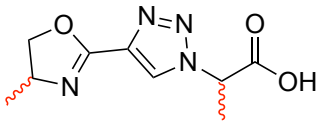
41



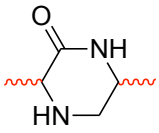
42



43

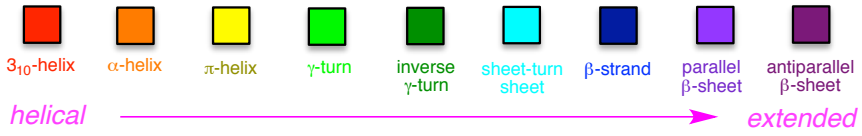


44

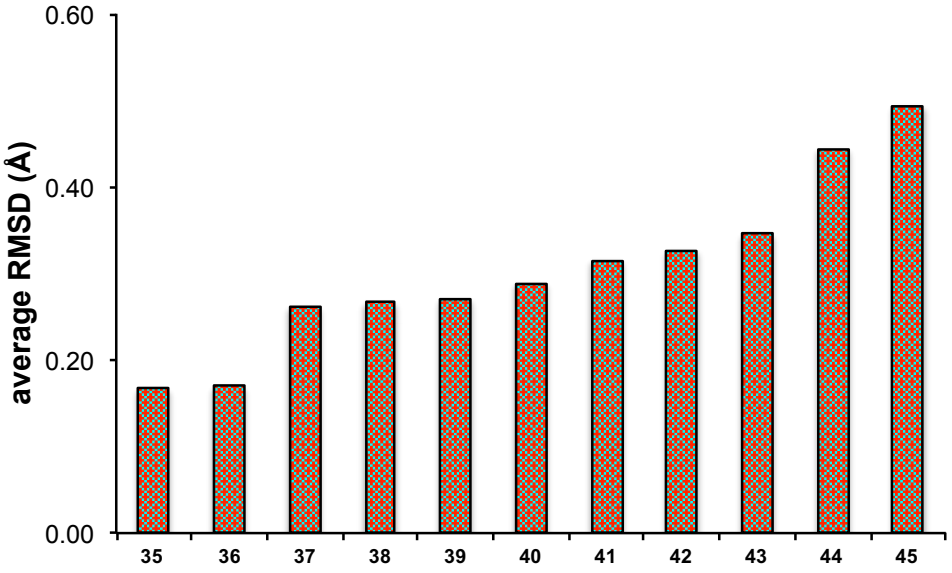


45

Ketopiperazine **35** and hydantoin **36** were the most promising dipeptide mimics in general; both have a bias towards extended structures; they fit better than the triazolamers core in **42**. While **38** is an inferior universal mimic, its bias towards helical structures is notable, especially since nearly all the mimics **A**, **1**, **2** and **24-29** prefer extended conformations. Chemotypes **35** and **36** are reasonable helical mimics, with a bias towards looser, π-helical forms, while tightly wound, 3_{10}-helical, conformations are preferred by **38**.



a



b

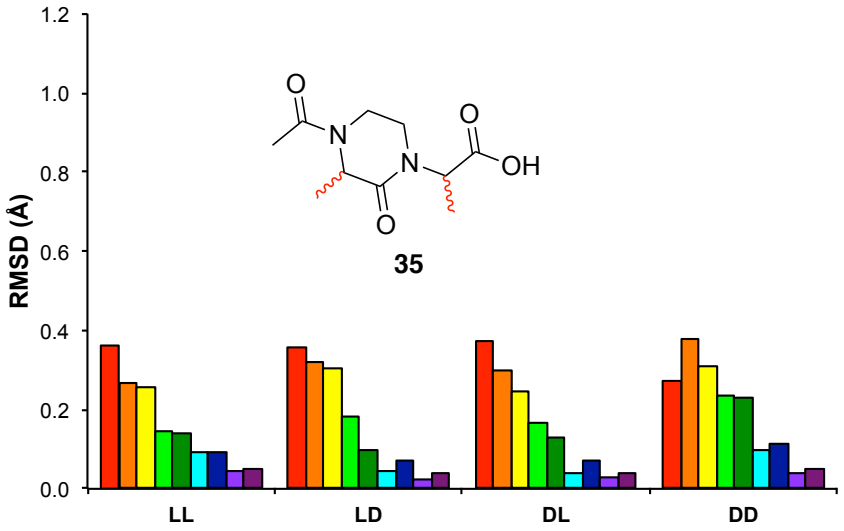
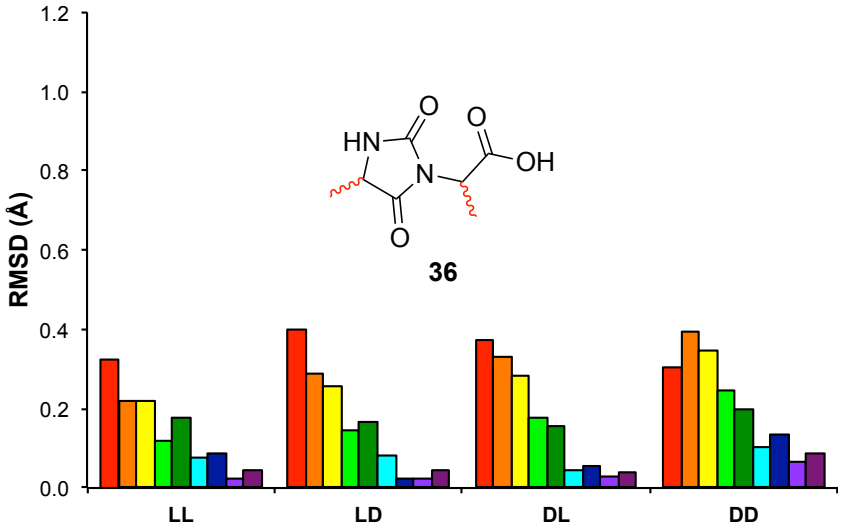


Figure E.1. (**a**) Average RMSD for overlays on all ideal secondary structures (except for **17** and **19** on turns and β-strands). (**b**) and (**c**) Overlays for the best general dipeptide

mimics in the series, *ie* **35** and **36**. (**d**) Overlays for the triazolamers core, **42**. (**e**) Overlays for a good dipeptide to use in helical mimics, **38**. Overlays for all other dipeptide mimics (**f** – **l**) represent **37**, **39**, **40**, **41**, **43**, **44**, **45**.

c



d

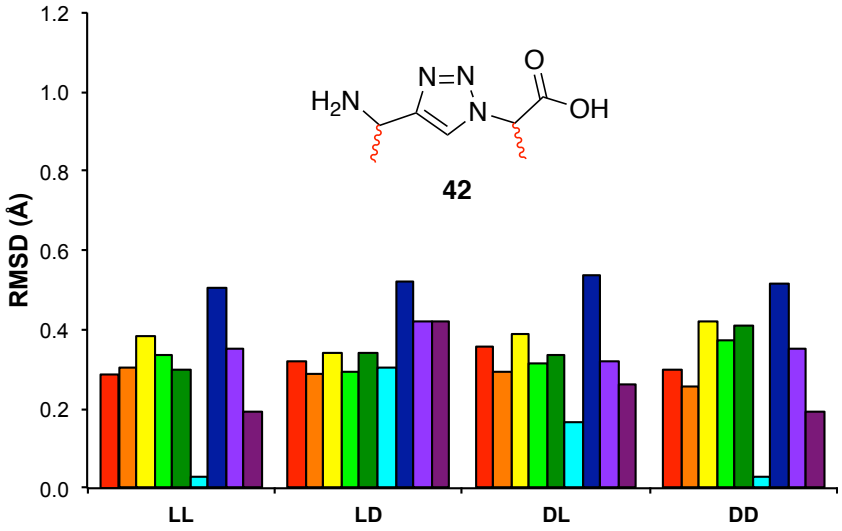
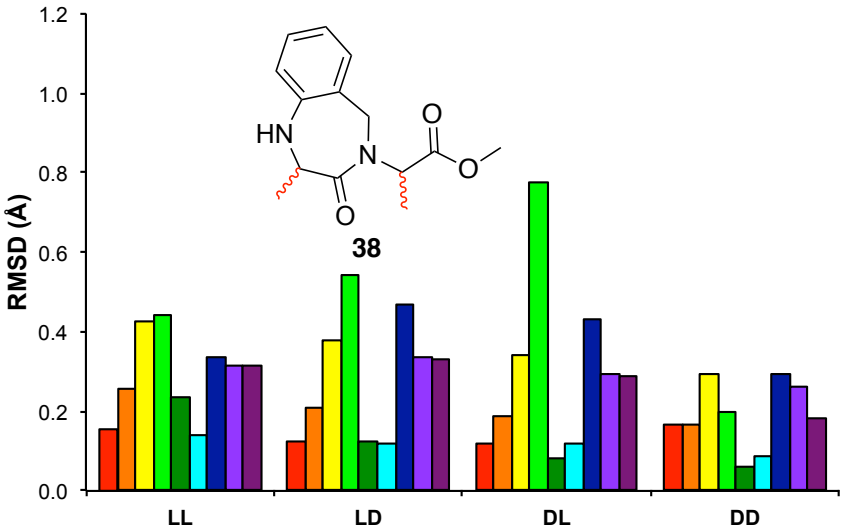


Figure E.1 Continued.

e



f

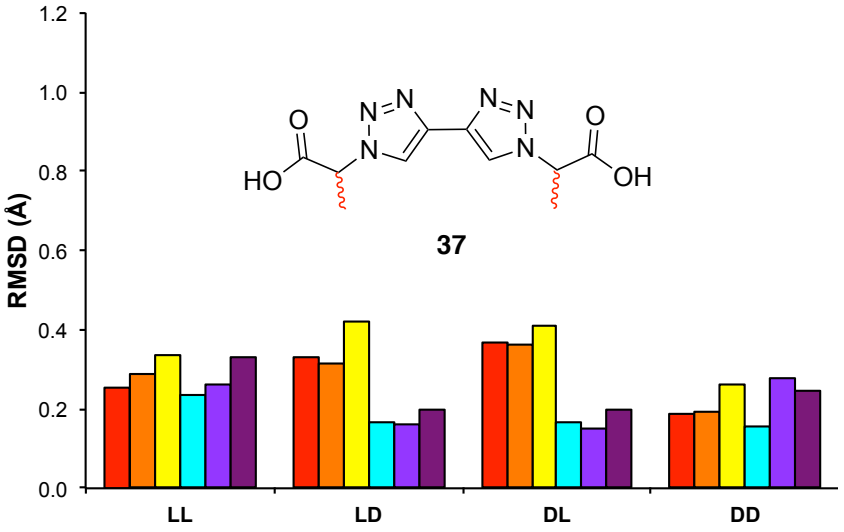
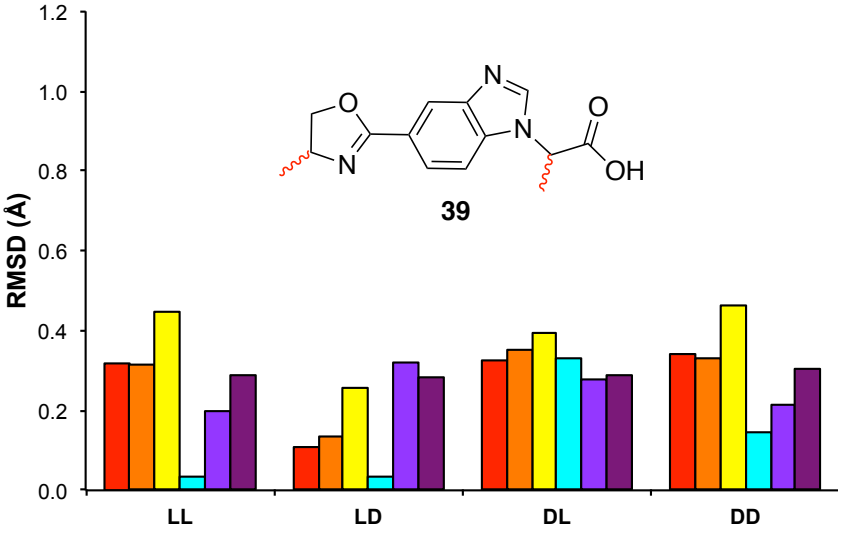


Figure E.1 Continued.

g



h

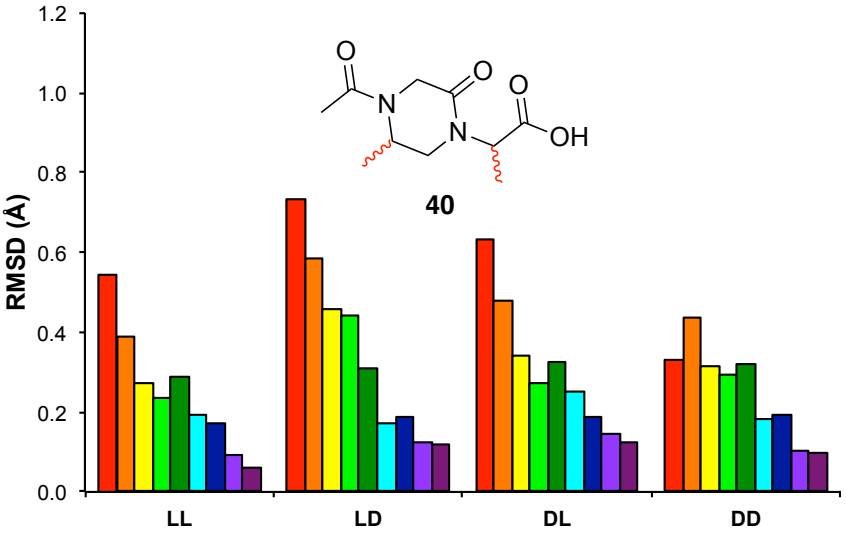
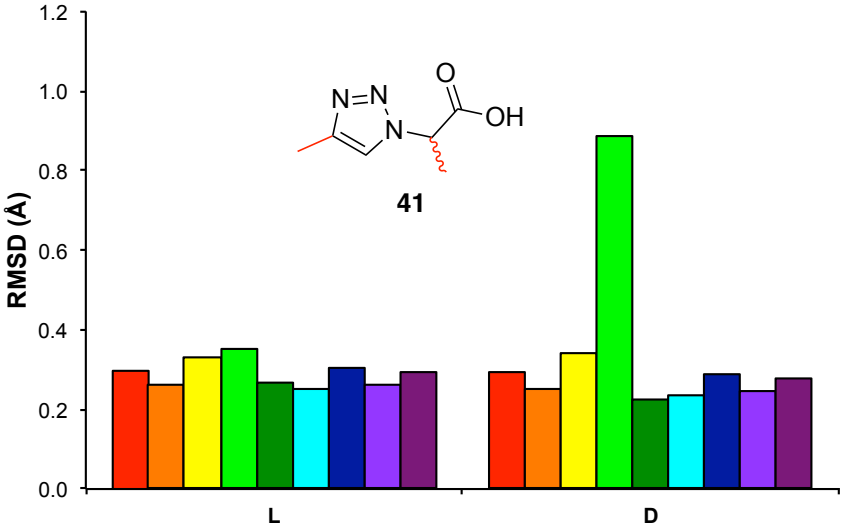


Figure E.1 Continued.

i



j

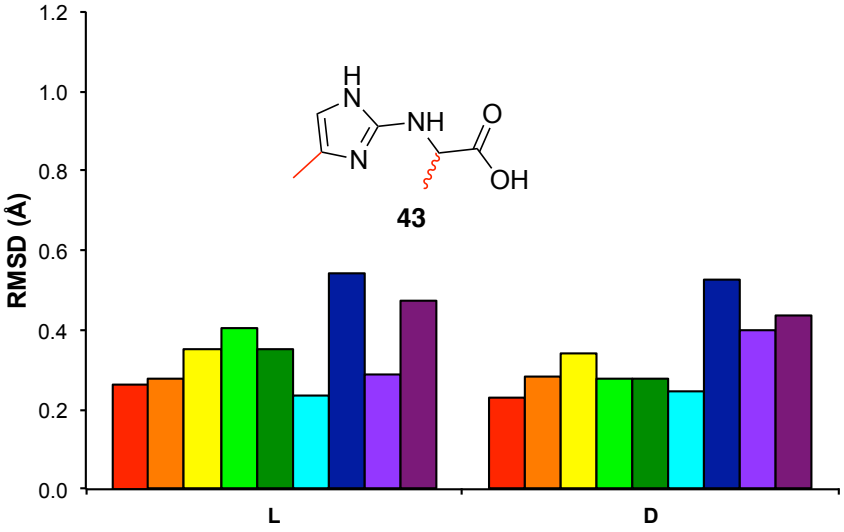
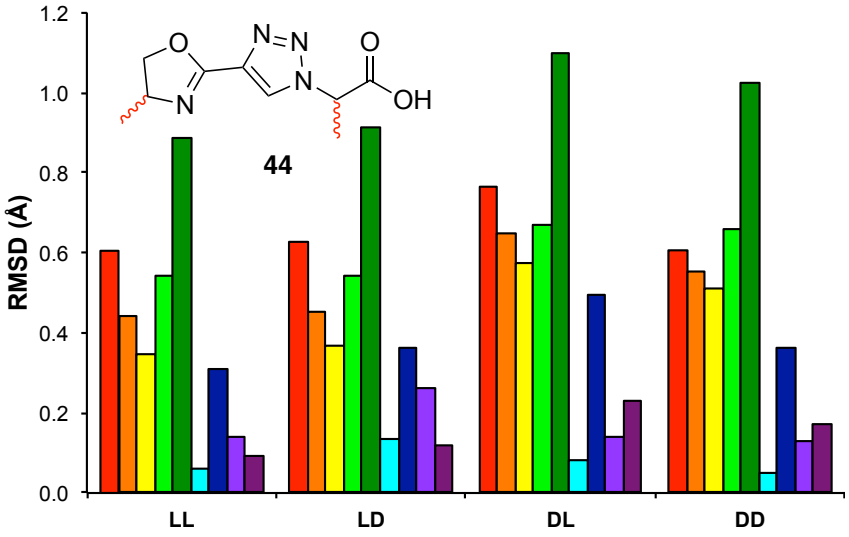


Figure E.1 Continued.

k



l

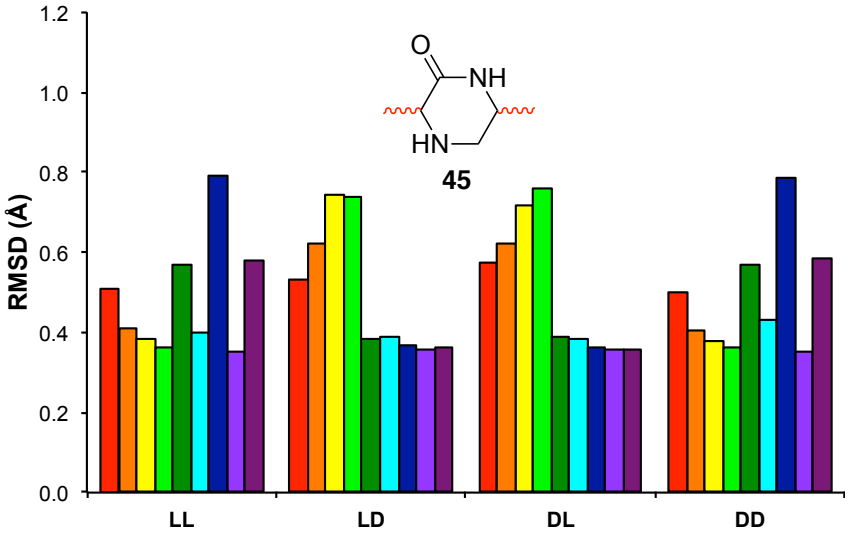


Figure E.1 Continued.

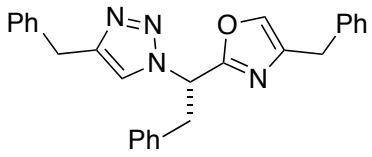
Table E.1 Number of conformers and clusters generated from QMD (< 3 kcal/mol from lowest energy).

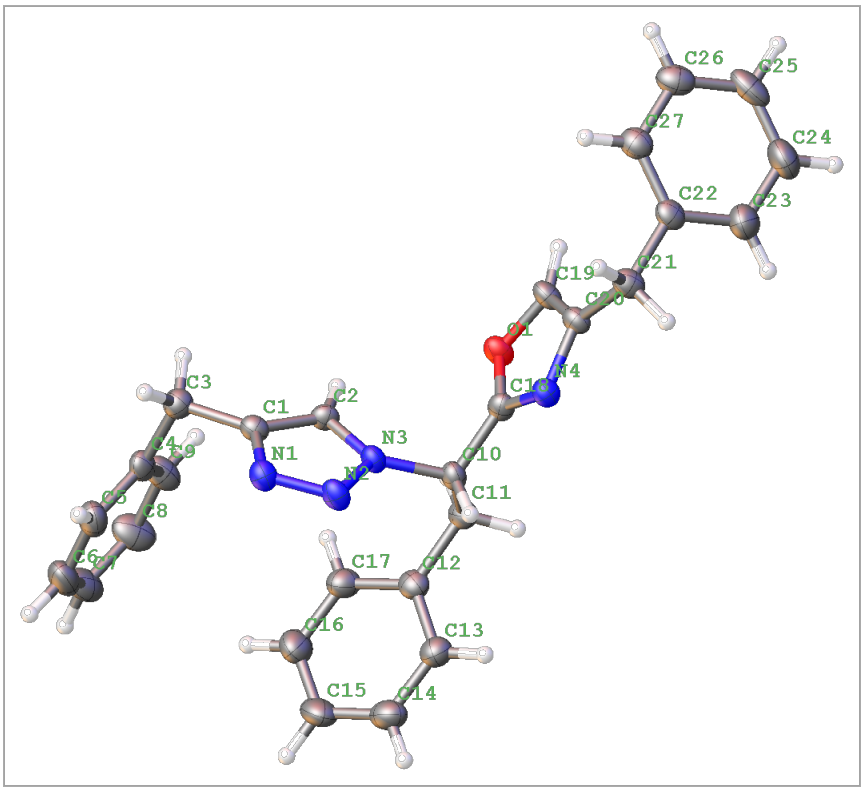
mimic	configuration	number of conformers	number of clusters
A	LLL	469	6
	LLD	782	6
	LDL	577	7
	DLL	636	8
	LDD	647	9
	DLD	820	12
	DDL	505	3
	DDD	310	3
1	LLL	1500	8
	LLD	1500	9
	LDL	1500	8
	DLL	1500	8
	LDD	1500	9
	DLD	1500	8
	DDL	1500	9
	DDD	1500	8
2	L	1500	3
	D	1500	3
24	LLL	1180	17
	LLD	1178	19
	LDL	1229	17
	DLL	1227	15
	LDD	1230	18
	DLD	1285	35
	DDL	1202	16
	DDD	1275	18

Table E.1 Continued.

25	LL	1476	19
	DL	1483	18
	LD	1485	16
	DD	1478	18
26	LLL	1500	22
	LLD	1500	21
	LDL	1500	23
	DLL	1500	22
	LDD	1500	21
	DLD	1500	22
	DDL	1500	21
	DDD	1500	23
27	LLL	1500	11
	LLD	1500	19
	LDL	1499	21
	DLL	1500	19
	LDD	1500	17
	DLD	1498	20
	DDL	1499	16
	DDD	1500	19
28	LLL	1499	32
	LLD	1498	38
	LDL	1499	36
	DLL	1499	33
	LDD	1498	37
	DLD	1497	36
	DDL	1497	35
	DDD	1498	33
29	L	1500	2
	D	1500	3

Crystal Structure Determination of ***L-2fff***





Crystal Data for $C_{27}H_{24}N_4O$ (*M* =420.50): orthorhombic, space group $P2_12_12_1$ (no. 19), *a* = 5.4354(6) Å, *b* = 12.4749(13) Å, *c* = 32.922(3) Å, *V* = 2232.3(4) Å^3, *Z* = 4, *T* = 110.0 K, μ(CuKα) = 0.616 mm^{-1}, *Dcalc* = 1.251 g/mm^3, 49719 reflections measured (5.368 ≤ 2Θ ≤ 128.728), 3669 unique (R_{int} = 0.0731, R_{sigma} = 0.0349) which were used in all calculations. The final R_1 was 0.0423 (I > 2σ(I)) and wR_2 was 0.1083 (all data).

APPENDIX F

EXPERIMENTAL PROCEDURES FOR CHAPTER VI

*Matching Results of **30** On Ideal Secondary Structures And Crystal Structure*

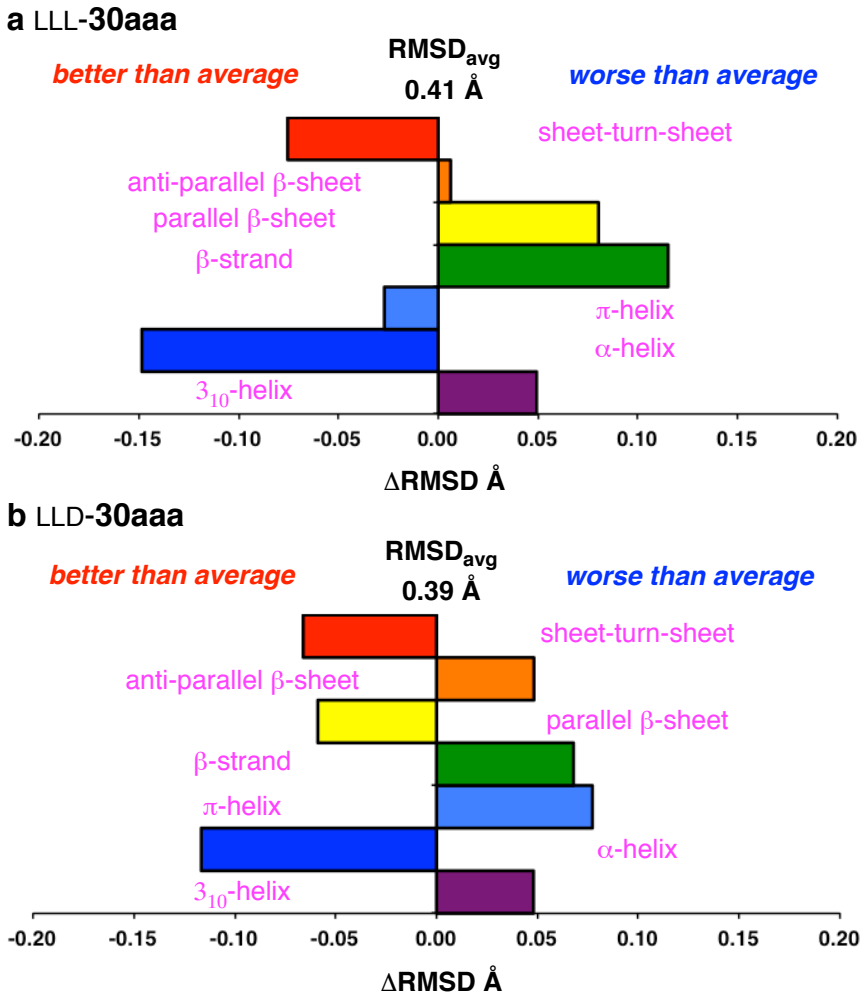


Figure F.1. Predicted mimicry of secondary structures for all diastereomers of **30aaa**

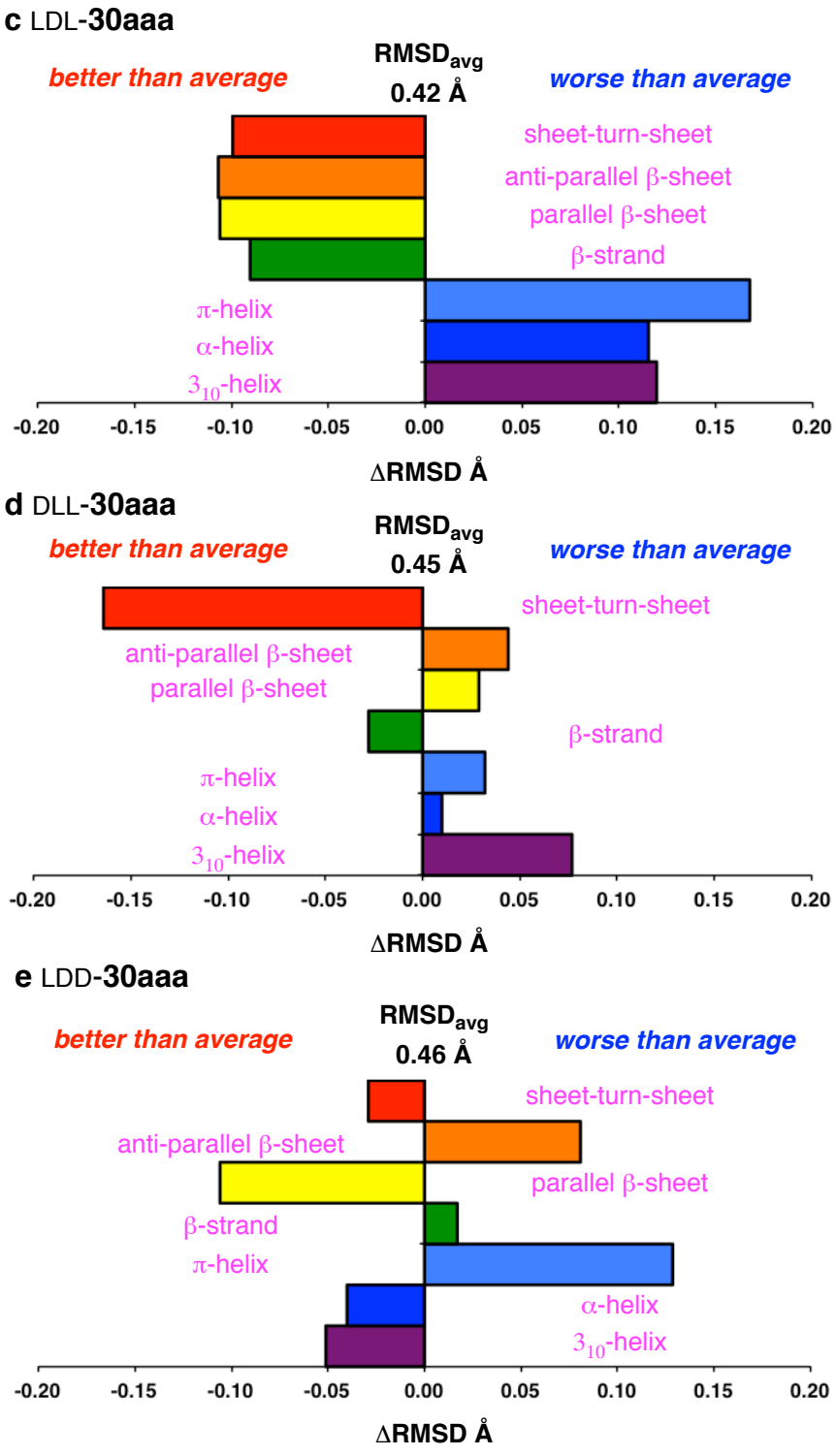
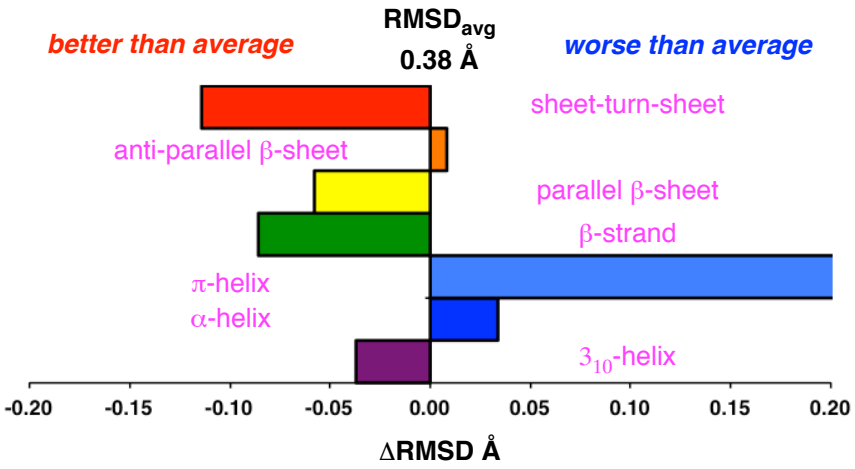
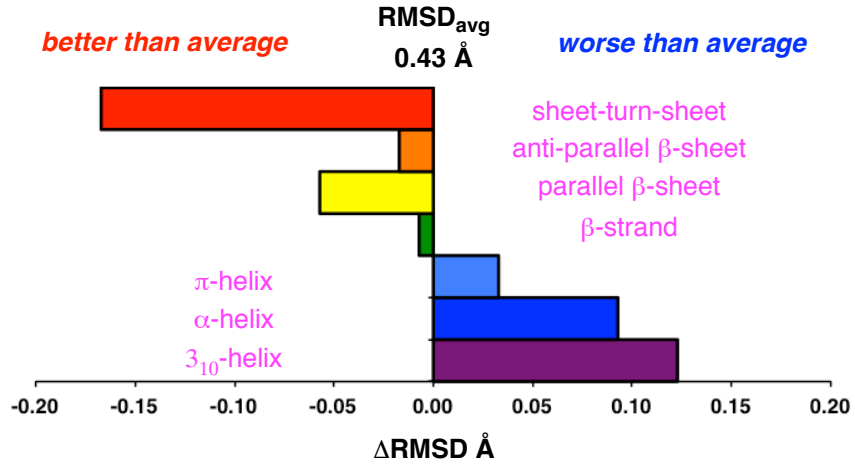


Figure F.1. Continued.

f DLD-**30aaa**



g DDL-**30aaa**



h DDD-**30aaa**

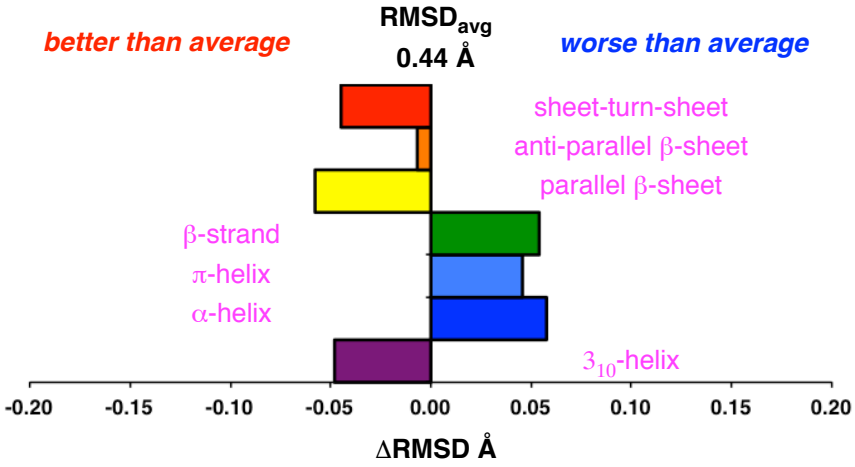
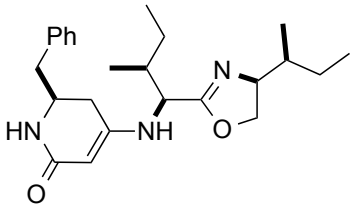
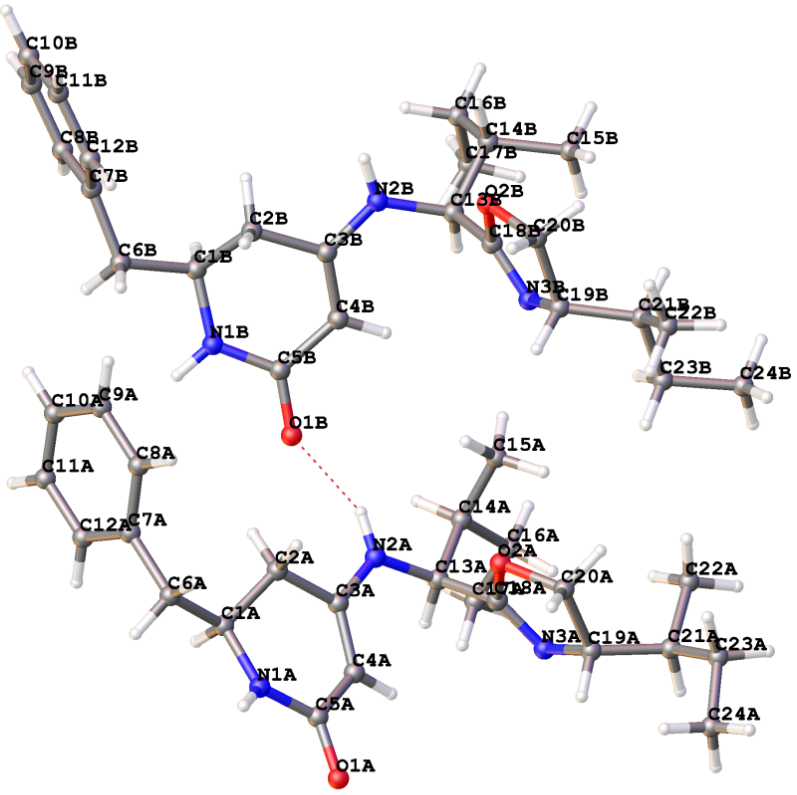


Figure F.1. Continued.

*X-Ray of LLL-**30-fii***





*Crystal Structure Determination Of LLL-**30fii***

Crystal Data for $C_{48}H_{71}N_6O_4$ (*M* =796.10): monoclinic, space group $P2_1$ (no. 4), *a* = 12.8948(11) Å, *b* = 9.7854(8) Å, *c* = 18.3889(16) Å, *β* = 94.667(6)°, *V* = 2312.6(3) $Å^3$, *Z* = 2, *T* = 110 K, μ(CuKα) = 0.572 mm^{-1}, *Dcalc* = 1.143 g/mm^3, 24722 reflections measured

(8.074 ≤ 2Θ ≤ 127.854), 7347 unique (R_{int} = 0.1116) which were used in all calculations. The final R_1 was 0.0694 (I > 2σ(I)) and wR_2 was 0.1819 (all data).

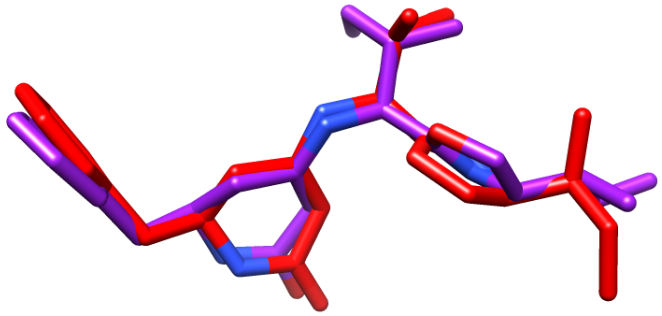


Figure F.2. Overlay of two different LLL-**30fii** conformers from crystal. Conformer 1 and 2 are shown in red and purple, respectively. RMSD = 0.28 Å based on the 6 *C*α and *C*β coordinates.

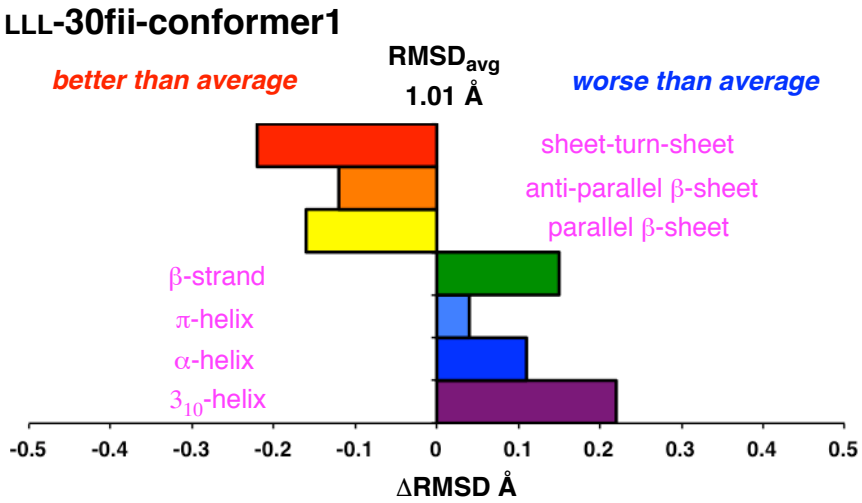


Figure F.3. Predicted mimicry of secondary structures for conformer 1 and 2 from crystal.

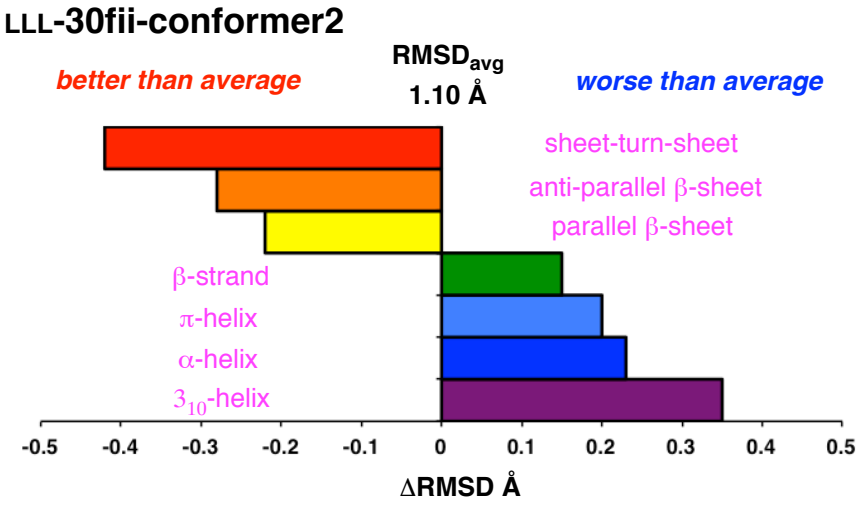
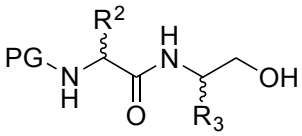


Figure F.3. Continued.

*General Procedure For The Preparation of C-terminal Dipeptide Amino Alcohol (**31**)*



31

Table F.1. Summary of dipeptide amino alcohol **31**

	PG	R^2	R^3	Yield (%)
L,D-**31vv**	Fmoc	L-Val	D-Val	87
L,L-**31ii**	Fmoc	L-Ile	L-Ile	90
L,D-**31ff**	Cbz	L-Phe	D-Phe	92
D,L-**31t'd'**	Fmoc	D-Thr(Bn)	L-Asp(^{t}Bu)	81
D,L-**31pv**	Fmoc	D-Pro	L-Val	93
D,L-**31pw**	Fmoc	D-Pro	L-Trp	80

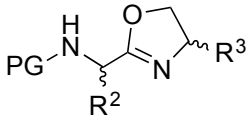
Table F.1. Continued.

L,L-**31py'**	**Fmoc**	**L-Pro**	**L-Tyr(^{t}Bu)**	**91**
D,L-**31d'k'**	Fmoc	D-Asp(^{t}Bu)	L-Lys(Cbz)	87
D,L-**31y's'**	Fmoc	D-Tyr(^{t}Bu)	L-Ser(Bn)	82
L,L-**31ae'**	Fmoc	L-Ala	L-Glu(^{t}Bu)	96
L,D-**31d't'**	Fmoc	L-Asp(^{t}Bu)	D-Thr(Bn)	71

A representative synthesis of the L,D-**31vv** is described as follows:

Unless the amino alcohols were commercially available, they are prepared from corresponding amino acids as previously described.[197] A solution of Fmoc-L-valine (4.91 g, 14.46 mmol) in dichloromethane (150 mL) was cooled to 0 ºC in an ice bath, D-valinol (1.79 g, 17.35 mmol), HOBt (2.93 g, 21.69 mmol) and EDCI (4.16 g, 21.69 mmol) were added in sequence. The resulting solution was stirred at 0 ºC for 15 min, then *N*-methyl morpholine (2.70 mL, 24.58 mmol) was added. The ice bath was removed and the reaction was stirred at ambient temperature for 4 – 5 h until the disappearance of starting material as monitored by TLC. The resulting solution was washed successively with cold 5 % $KHSO_4$ aqueous solution (3 × 50 mL), saturated $NaHCO_3$ solution (3 × 50 mL), and brine (75 mL). The combined organic layer was dried over $MgSO_4$ and filtered. The solvent was removed under vacuum and the resulting residue was crystallized using CH_2Cl_2-hexanes (1:4) to afford 5.35 g (87 %) of product as white solid. The compound was used in next step without further purification.

*General Procedure for the Preparation of N-protected Oxazoline (**32**)*

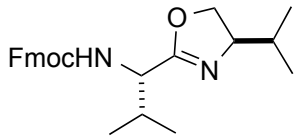


32

A representative synthesis of the L,D-**32vv** is described as follows:

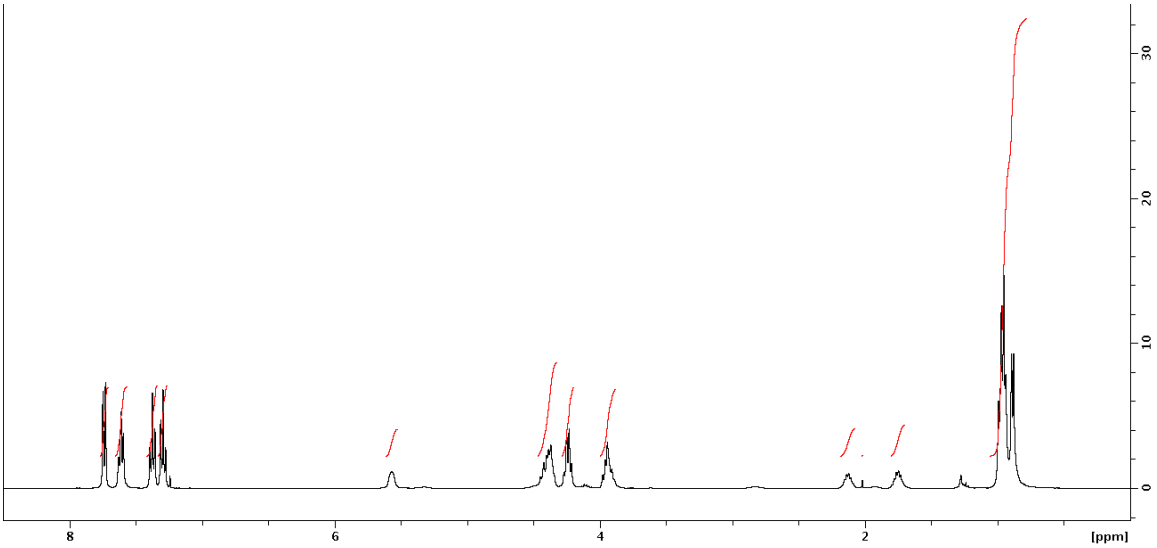
Compound L,D-**31vv** (0.25 g, 0.60 mmol) was dissolved in dry CH_2Cl_2 (40 mL) and cooled down using an ice bath. 4-Dimethylaminopyridine (7.33 mg, 0.06 mmol) and freshly distilled diisopropylethylamine (0.40 mL, 2.10 mmol) were added in sequence. The resulting clear solution was stirred at 0 ºC for 5min then a solution of methanesulfonyl chloride (0.06 mL, 0.72 mmol) in dry CH_2Cl_2 (50 mL) was added dropwise. The ice bath was removed and the reaction mixture was stirred at 25 ºC for 1 h until completion, as determined by TLC. The reaction was quenched with saturated NH_4Cl solution (50 mL) and the two phases were separated. The aqueous phase was extracted with CH_2Cl_2 (3 × 50 mL) and the combined organic layers were washed with brine (75 mL), then dried over anhydrous Na_2SO_4. Removing all the solvent gave crude product L,D-**32vv** which was chromatographed on silica gel with hexanes-EtOAc (6:4) then CH_2Cl_2-MeOH (98:2), affording the product (0.22 g, 92 %) as slightly yellow oil.

(9*H*-Fluoren-9-yl)methyl (*S*)-1-((*R*)-4-*iso*-propyl-4,5-dihydrooxazol-2-yl)-2-methylpropylcarbamate (L,D-**32vv**)

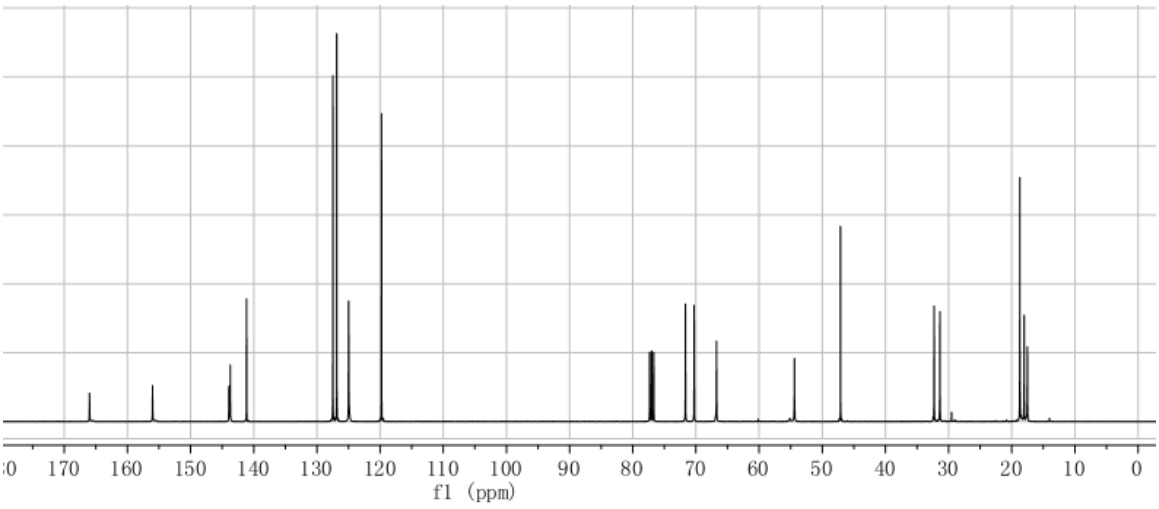


Data for L,D-32vv. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.70 (d, *J* = 7.4 Hz, 2H), 7.58 (t, *J* = 6.9 Hz, 2H), 7.34 (t, *J* = 7.3 Hz, 2H), 7.26 (t, *J* = 7.3 Hz, 2H), 5.66 (br, 1H), 4.36 (m, 3H), 4.20 (d, *J* = 7.0 Hz, 2H), 3.91 (m, 2H), 2.11 (m, 1H), 1.72 (m, 1H), 1.02 – 0.79 (m, 12H). ^{13}C NMR (100 MHz, $CDCl_3$) δ 166.2, 156.8, 155.9, 143.6, 141.0, 127.3, 126.7, 124.8, 119.6, 71.5, 70.1, 66.6, 54.3, 47.0, 32.1, 31.2, 18.6, 17.9. MS (ESI) m/z calcd for $C_{25}H_{31}N_2O_3^+$ 407.2329; found 407.2060 $(M+H)^+$.

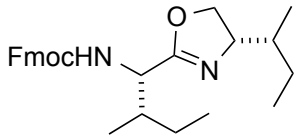
^{1}H NMR spectrum:



^{13}C NMR spectrum:

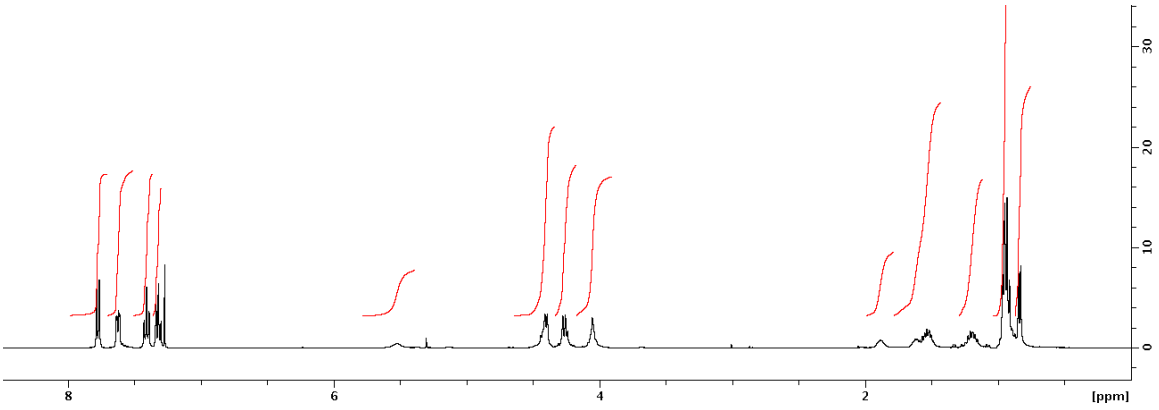


(9*H*-Fluoren-9-yl)methyl (1*S*,2*S*)-1-((*S*)-4-((*R*)-*sec*-butyl)-4,5-dihydrooxazol-2-yl)-2-methylbutylcarbamate (L,L-**32ii**)

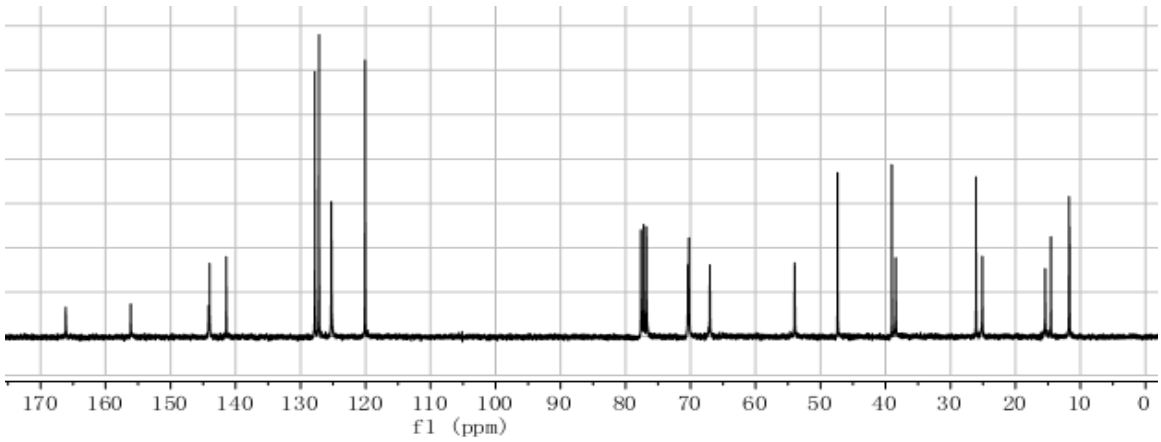


Data for L,L-32ii. White solid, 1.83 g (91%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.77 (d, *J* = 7.5 Hz, 2H), 7.63 (d, *J* = 7.4 Hz, 2H), 7.45 – 7.37 (m, 2H), 7.32 (m, 2H), 5.62 (br, 1H), 4.51 – 4.35 (m, 3H), 4.27 (m, 2H), 4.07 (br, 2H), 1.90 (br, 1H), 1.54 (m, 3H), 1.30 – 1.08 (m, 2H), 0.94 (dd, *J* = 12.1, 6.7 Hz, 9H), 0.84 (d, *J* = 6.7 Hz, 3H). ^{13}C NMR (100 MHz, $CDCl_3$) δ 166.0, 156.0, 143.9, 141.3, 127.6, 127.0, 125.2, 119.9, 70.3, 70.1, 66.9, 53.8, 47.3, 38.9, 38.3, 25.9, 24.9, 15.3, 14.4, 11.6, 11.5. HRMS (ESI) m/z calcd for $C_{27}H_{35}N_2O_3^+$ 435.2624; found 435.2651 $(M+H)^+$.

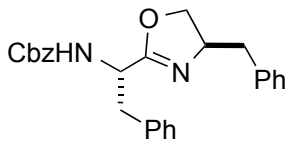
^{1}H NMR spectrum:



^{13}C NMR spectrum:

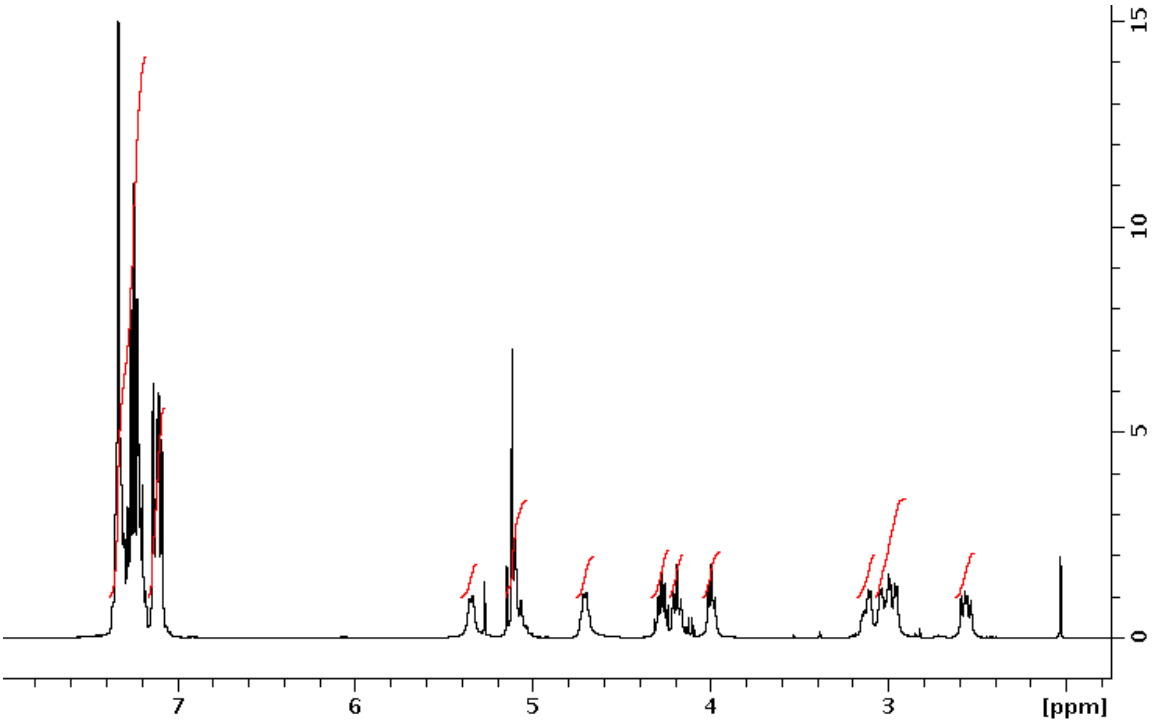

170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
f1 (ppm)

Benzyl (*S*)-1-((*S*)-4-benzyl-4,5-dihydrooxazol-2-yl)-2-phenylethylcarbamate (L,D-**32ff**)

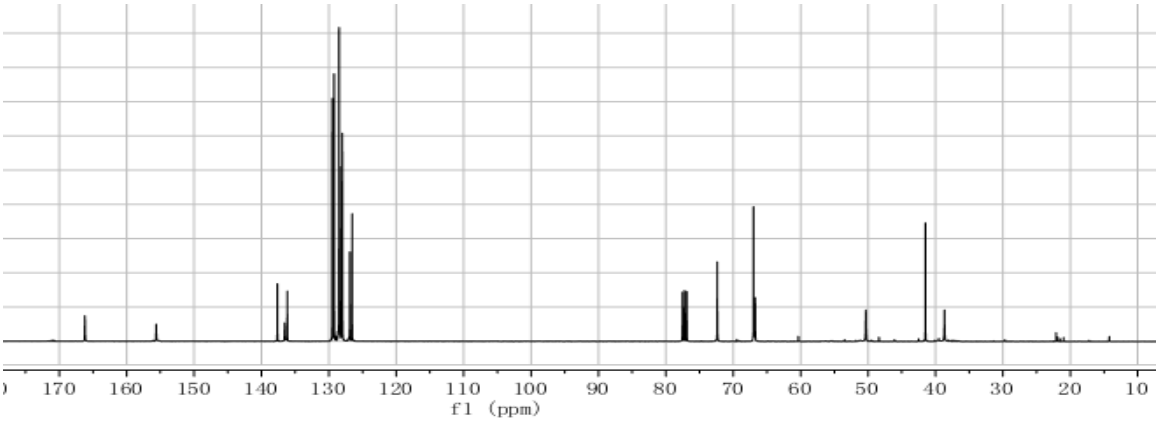


Data for L,D-32ff. White solid, 1.14 g (95%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.39-7.17 (m, 11H), 7.16-7.06 (m, 4H), 5.35 (br, 1H), 5.12 (m, 2H), 4.72 (br, 1H), 4.28 (m, 1H), 4.19 (m, 1H), 4.00 (m, 1H), 3.17-3.07 (m, 1H), 3.07-2.91 (m, 2H), 2.56 (m, 1H). ^{13}C NMR (100 MHz, $CDCl_3$) δ 166.2, 156.0, 137.7, 136.6, 136.2, 129.3, 129.2, 128.8, 128.7, 128.5, 128.2, 128.0, 127.1, 72.0, 67.1, 66.9, 50.1, 41.3, 38.4. HRMS (ESI) m/z calcd for $C_{26}H_{27}N_2O_3^+$ 415.2016; found 415.2031 $(M+H)^+$.

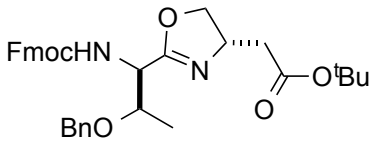
^{1}H NMR spectrum:



^{13}C NMR spectrum:

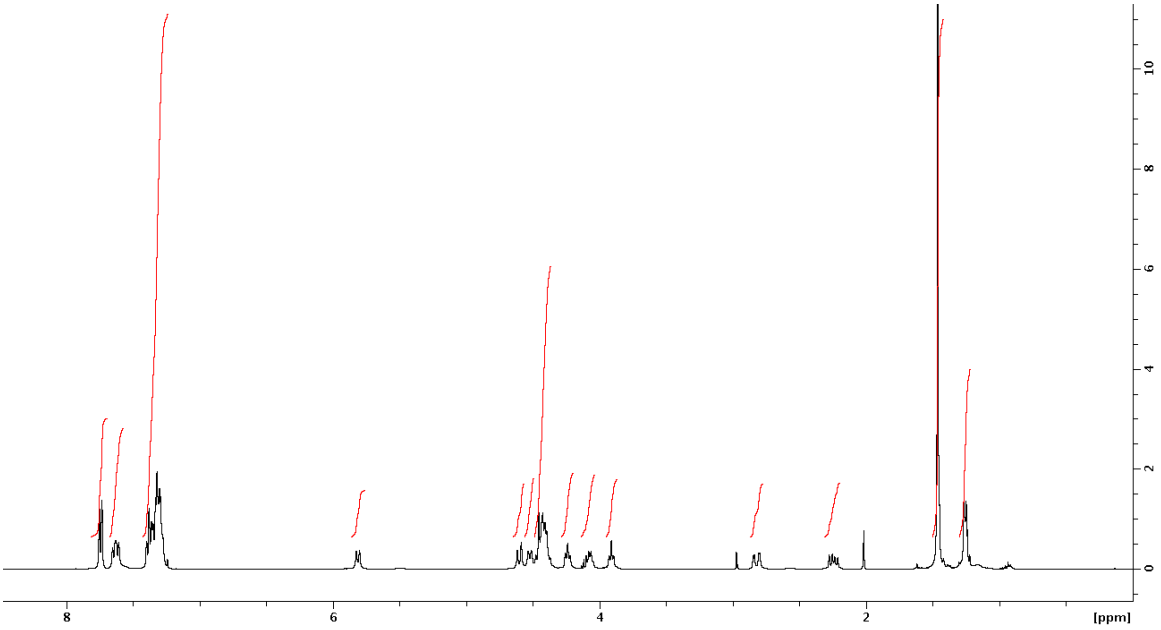


tert-Butyl 2-((*S*)-2-((1*R*,2*R*)-1-(((9*H*-fluoren-9-yl)methoxy)carbonyl)-2-(benzyloxy)propyl)-4,5-dihydrooxazol-4-yl)acetate (D,L-**32t'd'**)

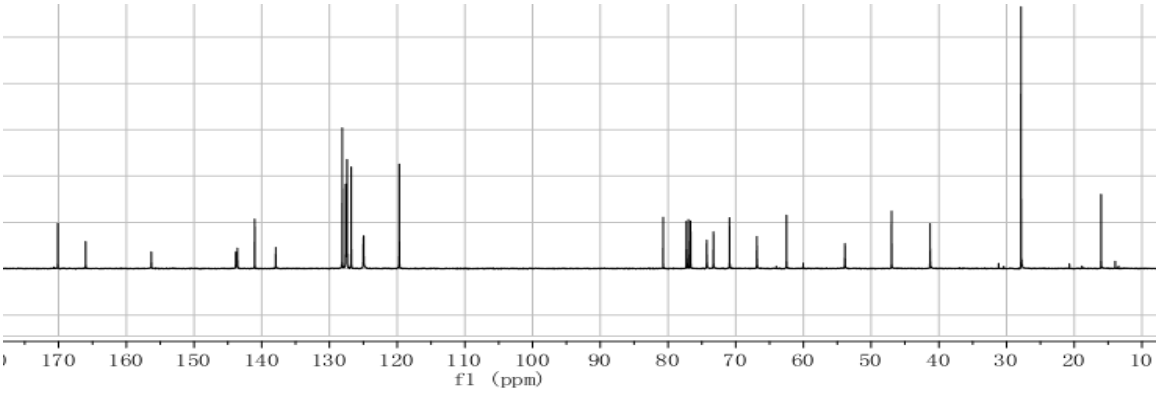


Data for D,L-32t'd'. White solid, 370 mg (81%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.74 (d, *J* = 7.5 Hz, 2H), 7.63 (m, 2H), 7.41 – 7.25 (m, 9H), 5.81 (d, *J* = 9.3 Hz, 1H), 4.60 (d, *J* = 11.8 Hz, 1H), 4.52 (d, *J* = 9.1 Hz, 1H), 4.49 – 4.35 (m, 5H), 4.24 (t, *J* = 7.1 Hz, 1H), 4.09 (m, 1H), 3.91 (t, *J* = 6.9 Hz, 1H), 2.82 (dd, *J* = 16.3, 3.5 Hz, 1H), 2.30 – 2.19 (m, 1H), 1.46 (s, 9H), 1.29 – 1.22 (m, 3H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 170.1, 166.0, 156.3, 143.6, 141.1, 137.9, 128.1, 127.6, 127.5, 127.4, 126.8, 125.0, 119.7, 80.7, 74.3, 73.3, 72.0, 70.9, 66.9, 62.5, 53.9, 47.0, 41.3, 27.9, 14.0; MS (ESI) m/z calcd for $C_{34}H_{39}N_2O_6^+$ 571.2803; found 571.2800 $(M+H)^+$.

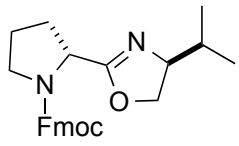
^{1}H NMR spectrum:



^{13}C NMR spectrum:



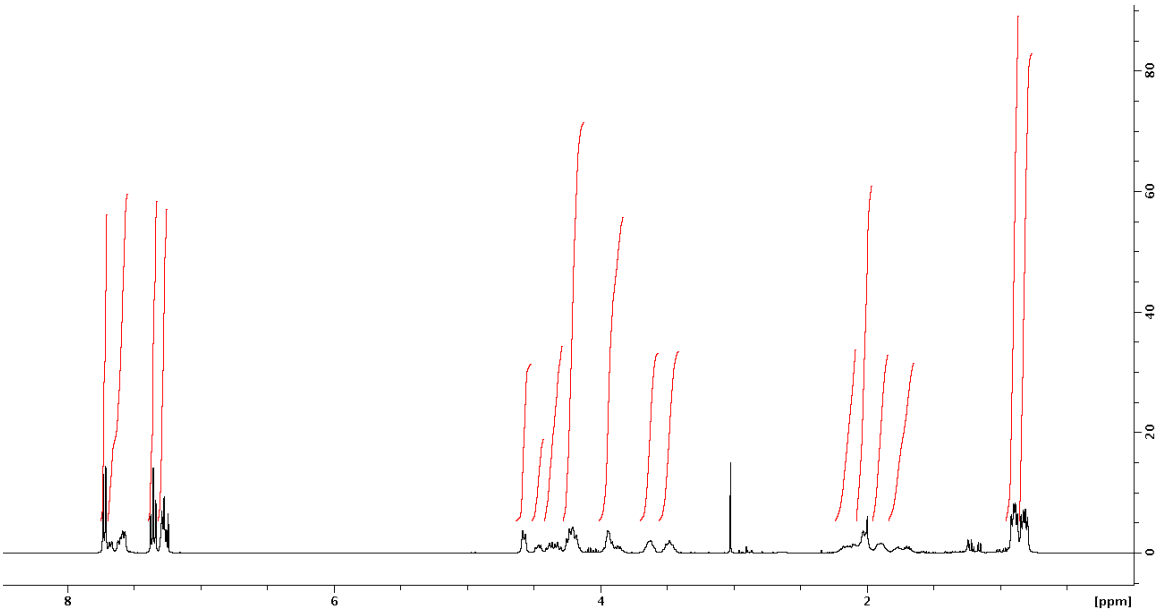
(*R*)-(9*H*-Fluoren-9-yl)methyl 2-((*S*)-4-*iso*-propyl-4,5-dihydrooxazol-2-yl)pyrrolidine-1-carboxylate (D,L-**32pv**)



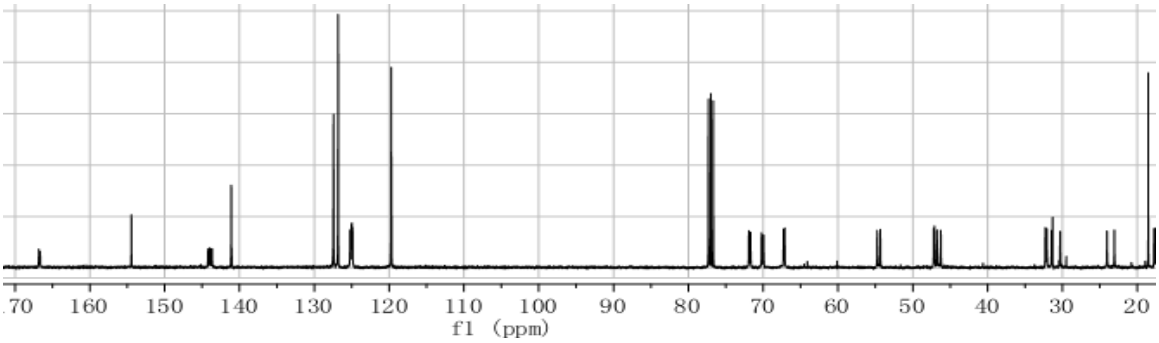
Data for D,L-32pv. Off-white crystal, 730 mg (54%) . ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.70 (d, *J* = 7.4 Hz, 2H), 7.59 (m, 2H), 7.35 (t, *J* = 7.3 Hz, 2H), 7.27 (m, 2H), 4.57 (d, *J* = 7.7

Hz, 1H), 4.46 (m, 1H), 4.35 (m, 1H), 4.21 (m, 2H), 3.98 – 3.81 (m, 2H), 3.69 – 3.57 (m, 1H), 3.55 – 3.41 (m, 1H), 2.22 – 2.06 (m, 1H), 2.06 – 1.96 (m, 2H), 1.96 – 1.83 (m, 1H), 1.83 – 1.63 (m, 1H), 0.95 – 0.86 (m, 3H), 0.82 (m, 3H). ^{13}C NMR (100 MHz, $CDCl_3$) δ 166.8, 154.4, 143.7, 141.1, 127.4, 126.8, 125.2, 125.0, 119.7, 71.9, 70.2, 70.0, 67.3, 55.0, 54.8, 47.5, 47.1, 46.3, 31.8, 30.8, 24.0, 23.1, 18.5. MS (ESI) m/z calcd for $C_{25}H_{29}N_2O_3^+$ 405.2173; found 405.2703 $(M+H)^+$

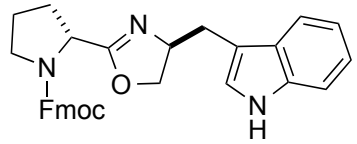
^{1}H NMR spectrum:



^{13}C NMR spectrum:

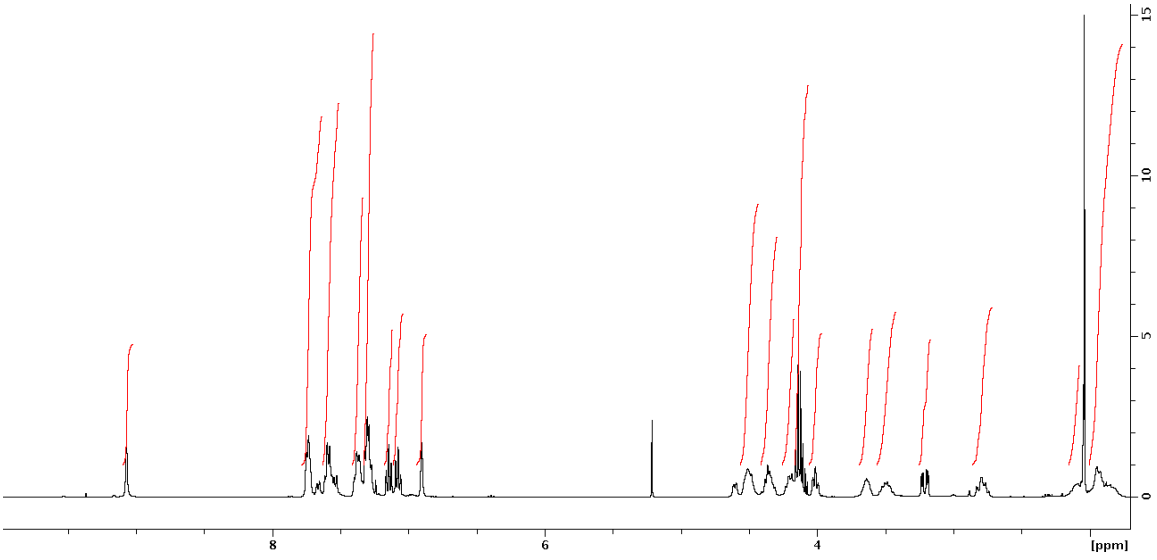


(*R*)-(9*H*-Fluoren-9-yl)methyl 2-((*S*)-4-((1*H*-indol-3-yl)methyl)-4,5-dihydrooxazol-2-yl)pyrrolidine-1-carboxylate (D,L-**32pw**)

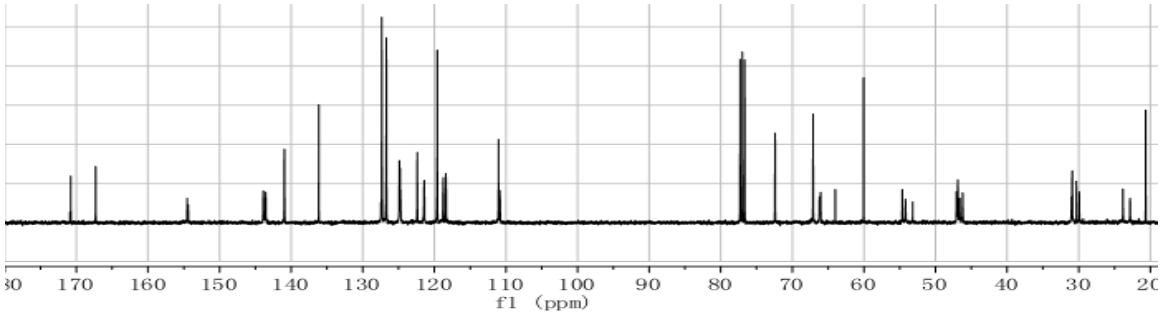


Data for D,L-**32pw**. White solid, 1.41 g (69%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.07 (s, 1H), 7.74 (m, 2H), 7.69 – 7.50 (m, 3H), 7.38 (m, 2H), 7.34 – 7.23 (m, 3H), 7.14 (t, *J* = 7.3 Hz, 1H), 7.07 (t, *J* = 7.3 Hz, 1H), 6.90 (s, 1H), 4.63 – 4.43 (m, 2H), 4.43 – 4.29 (m, 1H), 4.25 – 4.16 (m, 2H), 4.01 (t, *J* = 7.8 Hz, 1H), 3.63 (m, 1H), 3.56 – 3.43 (m, 1H), 3.21 (dd, *J* = 14.5, 4.6 Hz, 1H), 2.81 (m, 1H), 2.09 (d, *J* = 13.0 Hz, 1H), 2.01 – 1.85 (m, 4H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 170.8, 167.3, 154.5, 143.6, 140.9, 136.1, 127.3, 126.7, 124.9, 124.7, 122.4, 121.4, 119.6, 118.4, 111.9, 110.0, 72.4, 67.1, 66.2, 65.6, 64.0, 60.0, 54.9, 54.6, 47.1, 46.2, 33.2, 31.0, 20.7. HRMS (ESI) m/z calcd for $C_{31}H_{30}N_3O_3^+$ 492.2282; found 492.2292 $(M+H)^+$.

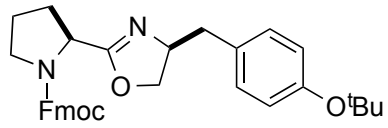
^{1}H NMR spectrum:



^{13}C NMR spectrum:

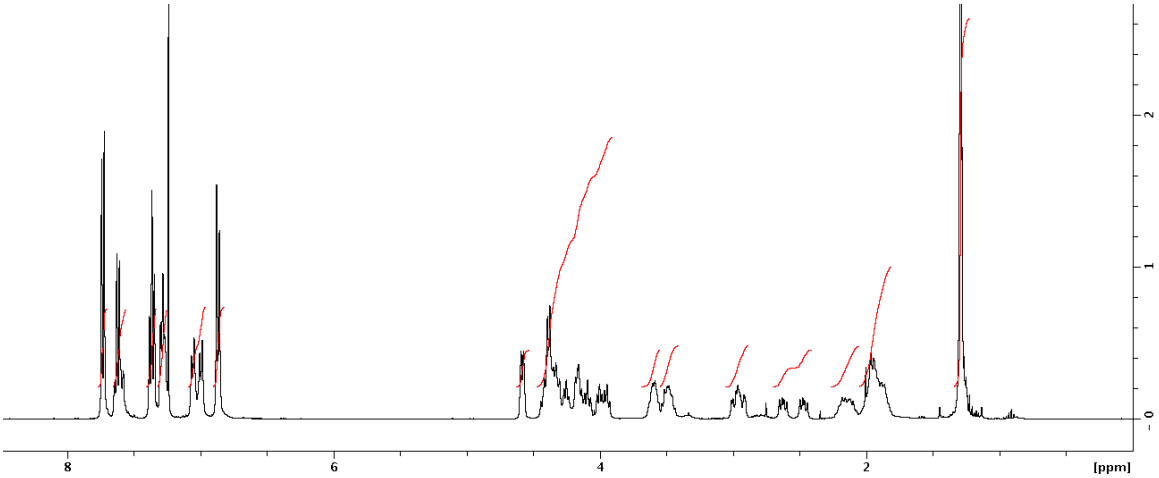


(*S*)-(9*H*-Fluoren-9-yl)methyl 2-((*S*)-4-(4-*tert*-butoxybenzyl)-4,5-dihydrooxazol-2-yl)pyrrolidine-1-carboxylate (L,L-**32py'**)

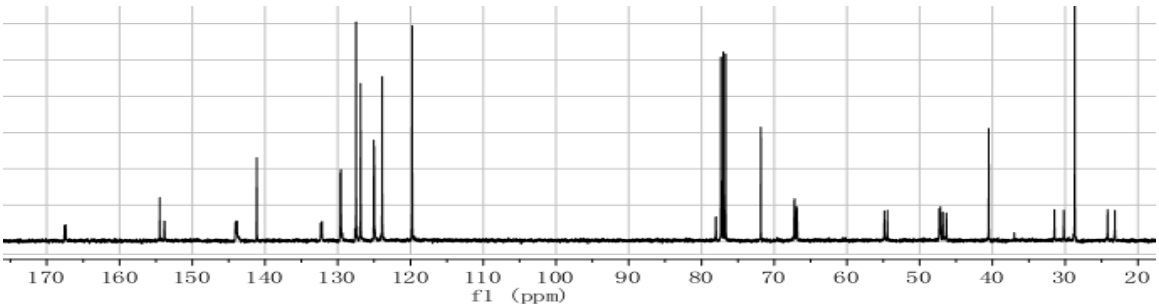


Data for L,L-32py'. White solid, 1.06 g (93%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.73 (d, *J* = 7.5 Hz, 2H), 7.66 – 7.56 (m, 2H), 7.36 (t, *J* = 7.4 Hz, 2H), 7.28 (t, *J* = 6.2 Hz, 2H), 7.06-6.96 (m, 3H), 6.87 (d, *J* = 8.3 Hz, 1H), 4.58 (d, *J* = 7.6 Hz, 1H), 4.46 – 3.91 (m, 6H), 3.58 (m, 1H), 3.49 (m, 1H), 3.03 – 2.89 (m, 1H), 2.55 (m, 1H), 2.26 – 2.05 (m, 1H), 1.93 (m, 3H), 1.29 (s, 9H). ^{13}C NMR (100 MHz, $CDCl_3$) δ 167.6, 154.5, 153.8, 144.1, 141.2, 132.4, 128.7, 128.4, 127.5, 126.9, 125.1, 119.8, 78.0, 71.8, 67.2, 66.9, 54.9, 54.7, 47.3, 47.0, 46.8, 46.5, 40.5, 32.2, 30.2, 28.7, 24.4, 24.2, 23.2. HRMS (ESI) m/z calcd for $C_{33}H_{37}N_2O_4^+$ 525.2748; found 525.2772 $(M+H)^+$

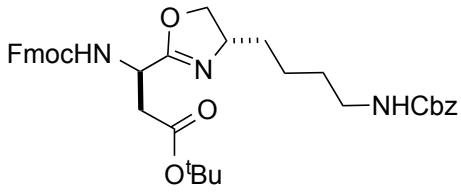
^{1}H NMR spectrum:



^{13}C NMR spectrum:



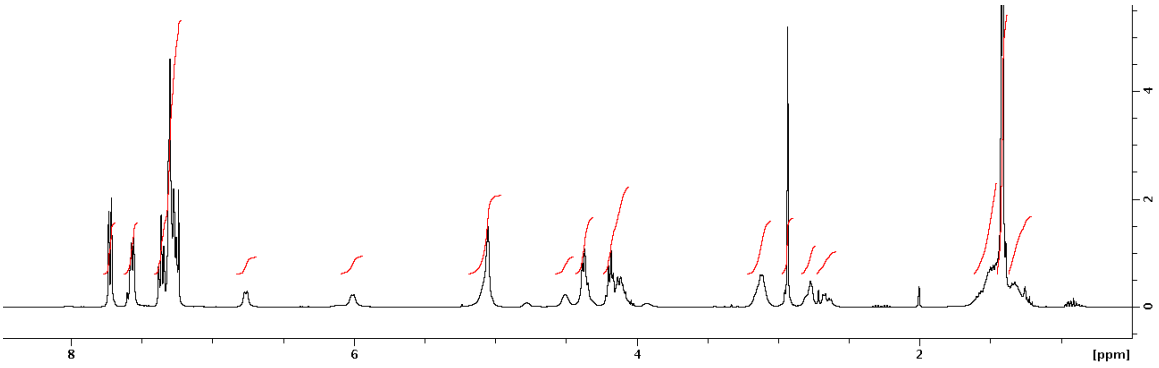
(*R*)-*tert*-Butyl 3-(((9*H*-fluoren-9-yl)methoxy)carbonyl)-3-((*S*)-4-(4-(benzyloxycarbonyl)butyl)-4,5-dihydrooxazol-2-yl)propanoate (D,L-**32d'k'**)



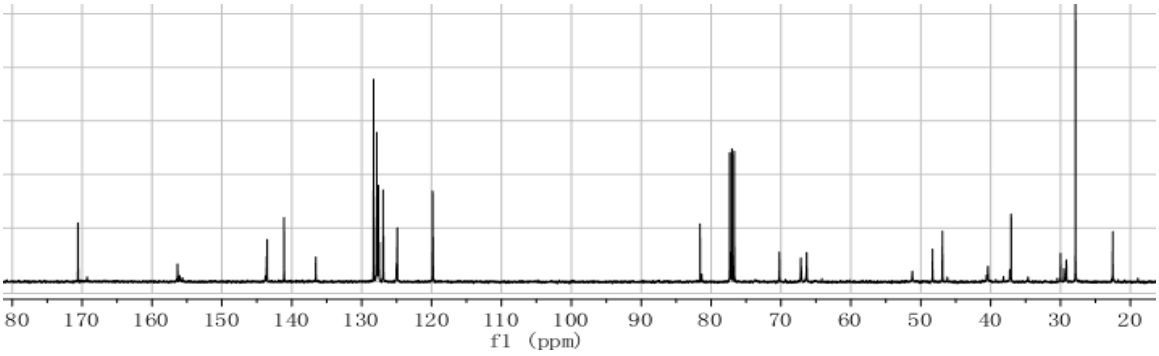
Data for D,L-32d'k'. Light yellow solid, 370 mg (80%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.71 (t, *J* = 7.5 Hz, 2H), 7.58 (t, *J* = 9.3 Hz, 2H), 7.36 (t, *J* = 7.4 Hz, 4H), 7.29 (dd, *J* = 15.9, 5.8 Hz, 5H), 6.76 (d, *J* = 7.4 Hz, 1H), 6.01 (d, *J* = 6.2 Hz, 1H), 5.05 (s, 2H), 4.50 (s, 1H), 4.37 (t, *J* = 7.9 Hz, 2H), 4.24 – 4.05 (m, 3H), 3.12 (d, *J* = 5.5 Hz, 2H), 2.94 (d, *J* = 8.1 Hz, 2H), 2.79 (d, *J* = 11.2 Hz, 1H), 2.73 – 2.58 (m, 1H), 1.59 – 1.42 (m, 3H), 1.41 (s,

9H), 1.36 – 1.19 (m, 2H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 170.6, 169.3, 156.4, 156.1, 143.6, 141.1, 136.6, 128.3, 127.8, 127.5, 126.9, 125.0, 124.9, 119.9, 119.8, 81.6, 70.2, 67.2, 66.3, 51.2, 48.3, 47.0, 40.4, 37.3, 30.0, 29.2, 22.5; MS (ESI) m/z calcd for $C_{37}H_{44}N_3O_7^+$ 642.32; found 642.24 $(M+H)^+$.

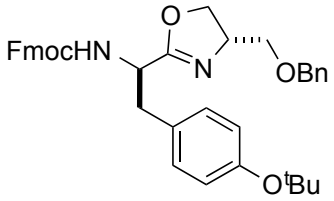
^{1}H NMR spectrum:



^{13}C NMR spectrum:

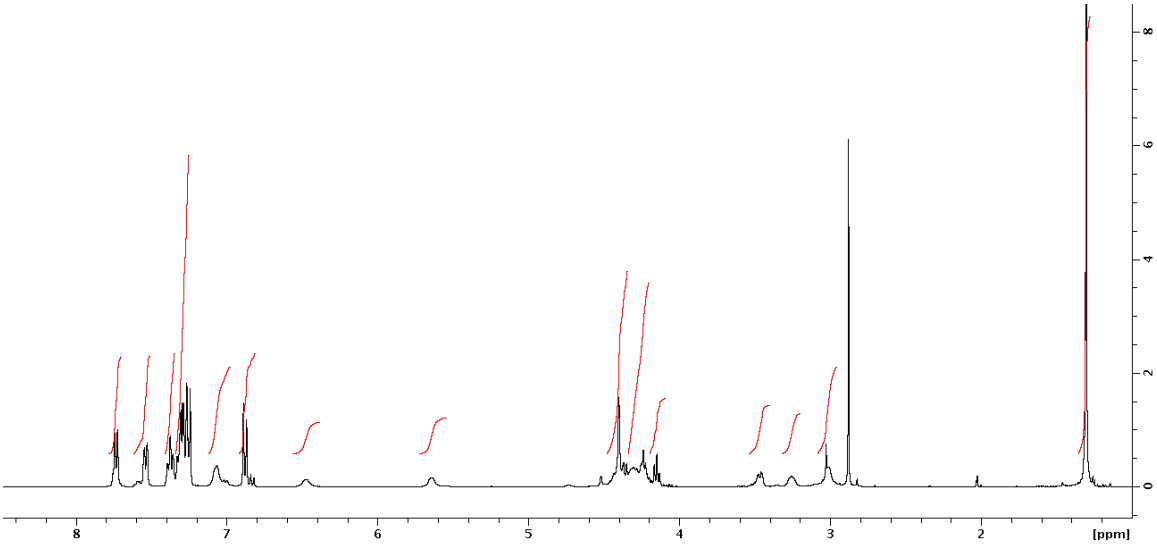


(9*H*-Fluoren-9-yl)methyl (*R*)-1-((*S*)-4-(benzyloxymethyl)-4,5-dihydrooxazol-2-yl)-2-(4-*tert*-butoxyphenyl)ethylcarbamate (D,L-**32y's'**)

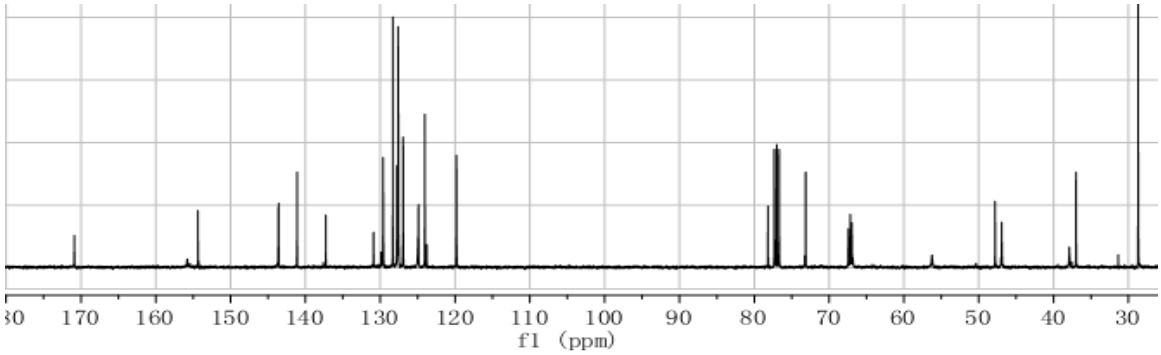


Data for D,L-**32y's'**. Yellow solid, 980 mg (98%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.74 (t, *J* = 6.1 Hz, 2H), 7.53 (d, *J* = 7.4 Hz, 2H), 7.36 (q, *J* = 7.5 Hz, 2H), 7.34 – 7.25 (m, 6H), 7.12 – 6.98 (m, 2H), 6.85 (d, *J* = 8.4 Hz, 2H), 6.47 (s, 1H), 5.64 (s, 1H), 4.47 – 4.34 (m, 4H), 4.34 – 4.19 (m, 4H), 4.19 – 4.12 (m, 1H), 3.53 – 3.43 (m, 1H), 3.26 (s, 1H), 3.05 – 2.97 (m, 2H), 1.30 (s, 9H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 170.9, 155.8, 154.3, 143.6, 141.1, 137.6, 137.3, 130.9, 129.6, 128.3, 127.8, 127.6, 127.5, 127.0, 124.9, 124.0, 123.8, 119.8, 78.2, 73.1, 67.4, 67.0, 56.3, 47.8, 37.0, 28.7; HRMS (ESI) m/z calcd for $C_{38}H_{41}N_2O_5^+$ 605.3010; found 605.3020 $(M+H)^+$.

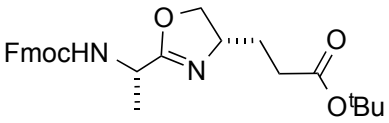
^{1}H NMR spectrum:



^{13}C NMR spectrum:

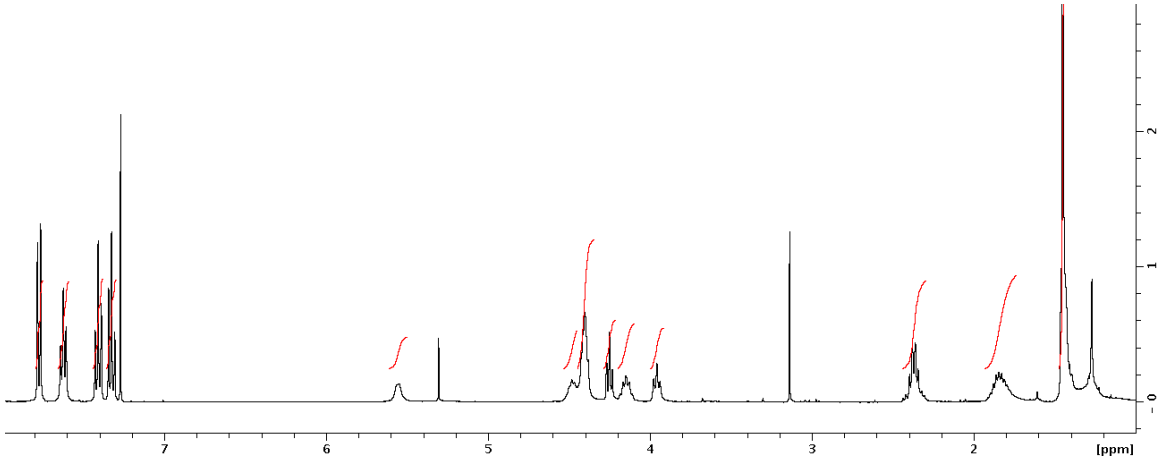


tert-Butyl 3-((*R*)-2-((*S*)-1-(((9*H*-fluoren-9-yl)methoxy)carbonyl)ethyl)-4,5-dihydrooxazol-4-yl)propanoate (L,L-**32ae'**)

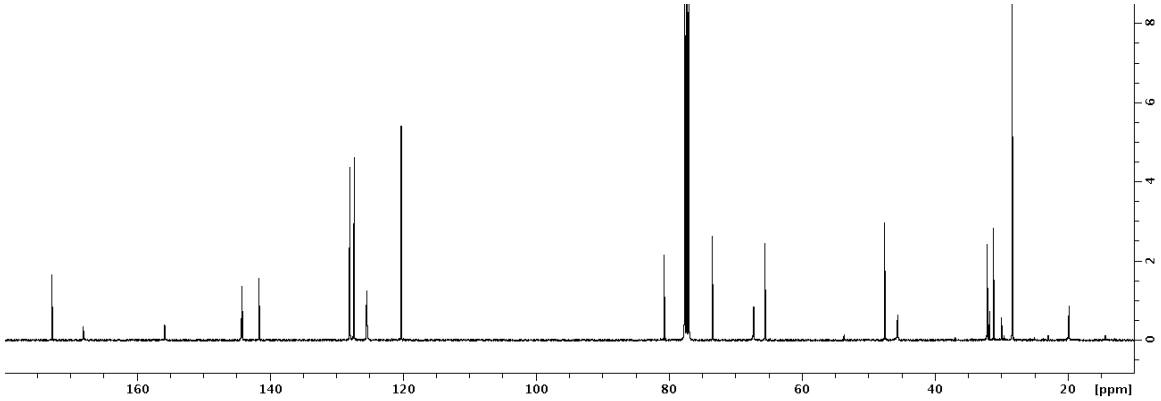


Data for L,L-32ae'. Yellow solid, 540 mg (98%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.77 (d, *J* = 7.3 Hz, 2H), 7.62 (t, *J* = 6.8 Hz, 2H), 7.40 (t, *J* = 7.3 Hz, 2H), 7.32 (t, *J* = 7.5 Hz, 2H), 5.56 (br, 1H), 4.48 (m, 1H), 4.40 (m, 3H), 4.15 (m, 1H), 3.96 (t, *J* = 8.1 Hz, 1H), 3.16 – 2.95 (m, 1H), 2.37 (m, 2H), 1.85 (m, 2H), 1.45 (m, 12H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 172.7, 168.0, 155.7, 144.1, 141.5, 127.9, 127.2, 125.4, 120.2, 80.6, 73.4, 67.2, 65.5, 47.4, 45.5, 32.0, 31.7, 31.1, 29.88, 28.3, 19.7. HRMS (ESI) m/z calcd for $C_{27}H_{33}N_2O_5^+$ 465.2384; found 465.2383 $(M+H)^+$.

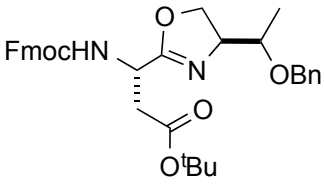
^{1}H NMR spectrum:



^{13}C NMR spectrum:



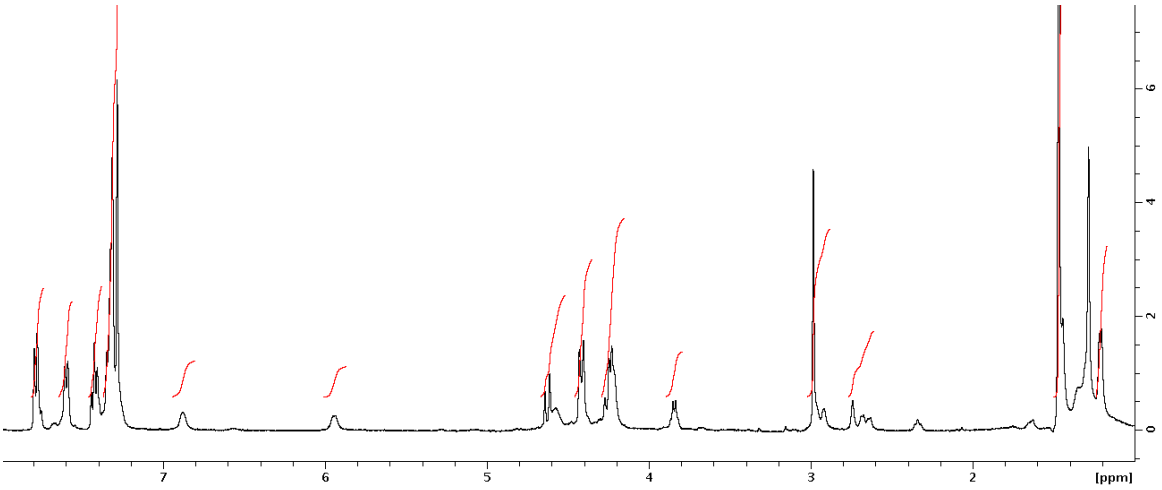
(*S*)-*tert*-Butyl 3-(((9*H*-fluoren-9-yl)methoxy)carbonyl)-3-((*S*)-4-((*S*)-1-(benzyloxy)ethyl)-4,5-dihydrooxazol-2-yl)propanoate (L,D-**32d't'**)



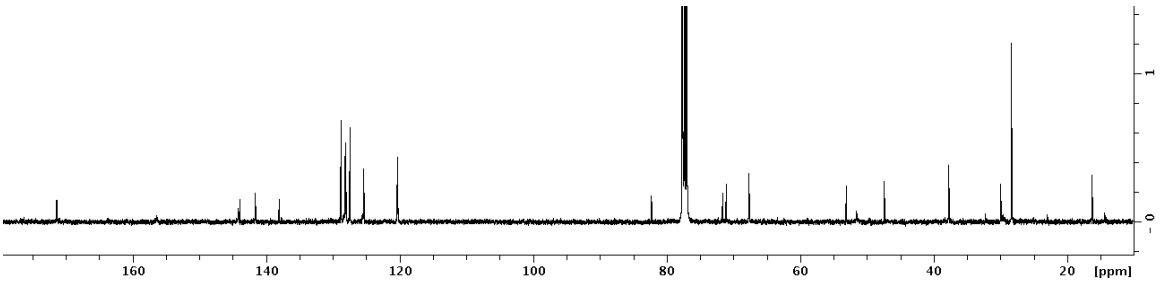
Data for L,D-32d't'. Yellow solid, 380 mg (84%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.79 (d, *J* = 7.4 Hz, 2H), 7.60 (d, *J* = 7.7 Hz, 2H), 7.43 (t, *J* = 7.4 Hz, 2H), 7.36 – 7.24 (m, 7H), 6.88 (br, 1H), 5.94, (br, 1H) 4.65-4.53 (m, 2H), 4.42 (d, *J* = 10.3 Hz, 2H), 4.28 – 4.17 (m, 3H), 3.84 (m, 1H), 3.00-2.90 (m, 3H), 2.77-2.61 (m, 1H), 1.47 (s, 9H), 1.21 (d, *J* = 5.6 Hz,

3H). ^{13}C NMR (100 MHz, $CDCl_3$) δ 171.3, 156.3, 144.1, 143.9, 141.6, 141.5, 138.0, 128.7, 128.1, 128.01, 127.9, 127.3, 125.3, 82.2, 71.5, 71.0, 67.6, 53.0, 47.4, 37.7, 29.9, 28.3, 16.2. HRMS (ESI) m/z calcd for $C_{34}H_{38}N_2O_6^+$ 571.2803; found 571.2816 $(M+H)^+$.

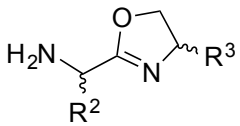
^{1}H NMR spectrum:



^{13}C NMR spectrum:



*General Procedure For The Preparation of Amino Oxazoline (**33**)*



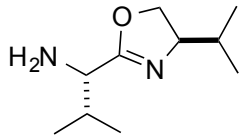
33

A representative synthesis of the L,D-**33vv** is described as follows:

To a solution of L,D-**32vv** (301.5 mg, 0.742 mmol) in MeCN (20 mL), diethylamine (0.8 mL, 7.42 mmol) was added in one portion at room temperature. The reaction was stirred

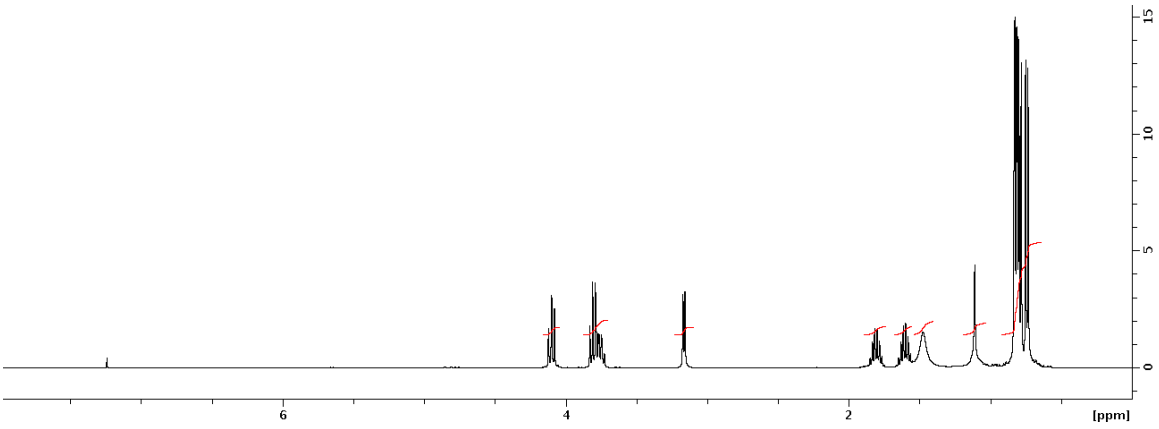
for 2 h and the solvent was removed under vacuum. Residual Et_2NH was azeotroped 3 times with CH_2Cl_2 (3 × 20 mL). The crude product was purified by flash chromatography (silica gel, hexane-EtOAc, 9:1 then dichloromethane-methanol, 20:1) to give the free amino-containing oxazoline compound L,D-**33vv** as light yellow oil (110 mg, 81 %).

(*S*)-1-((*R*)-4-*iso*-Propyl-4,5-dihydrooxazol-2-yl)-2-methylpropan-1-amine (L,D-**33vv**)

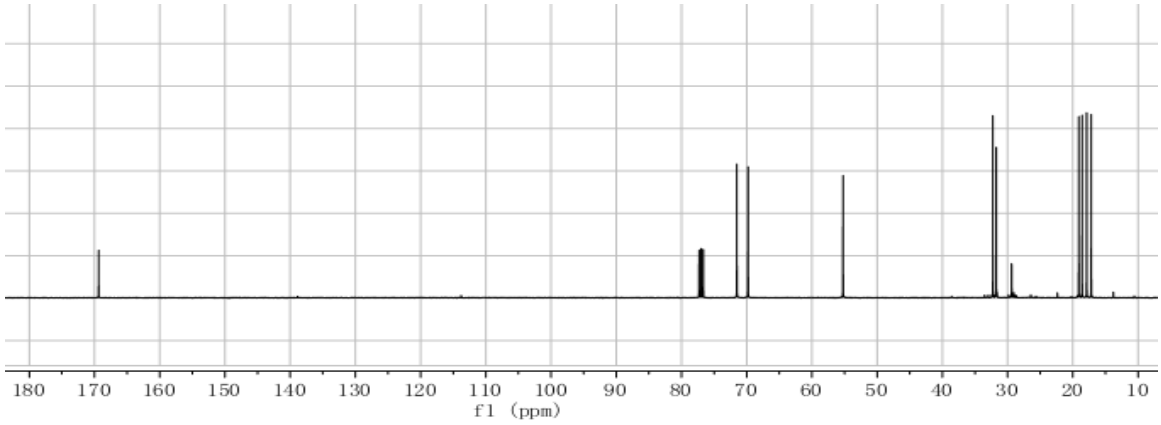


Data for L,D-33vv. ^{1}H NMR (400 MHz, $CDCl_3$) δ 4.10 (dd, *J* = 8.9, 7.7 Hz, 1H), 3.85 – 3.71 (m, 1H), 3.16 (d, *J* = 5.6 Hz, 1H), 1.86 – 1.75 (m, 1H), 1.67 – 1.54 (m, 1H), 1.47 (br, 2H), 1.10 (br, 1H), 0.88 – 0.69 (m, 12H). ^{13}C NMR (100 MHz, $CDCl_3$) δ 169.4, 71.5, 69.8, 55.2, 32.3, 31.8, 19.1, 18.6, 17.9, 17.2. HRMS (ESI) m/z calcd for $C_{10}H_{21}N_2O^+$ 185.1648; found 185.1661 $(M+H)^+$.

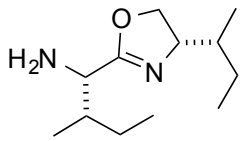
^{1}H NMR spectrum:



^{13}C NMR spectrum:

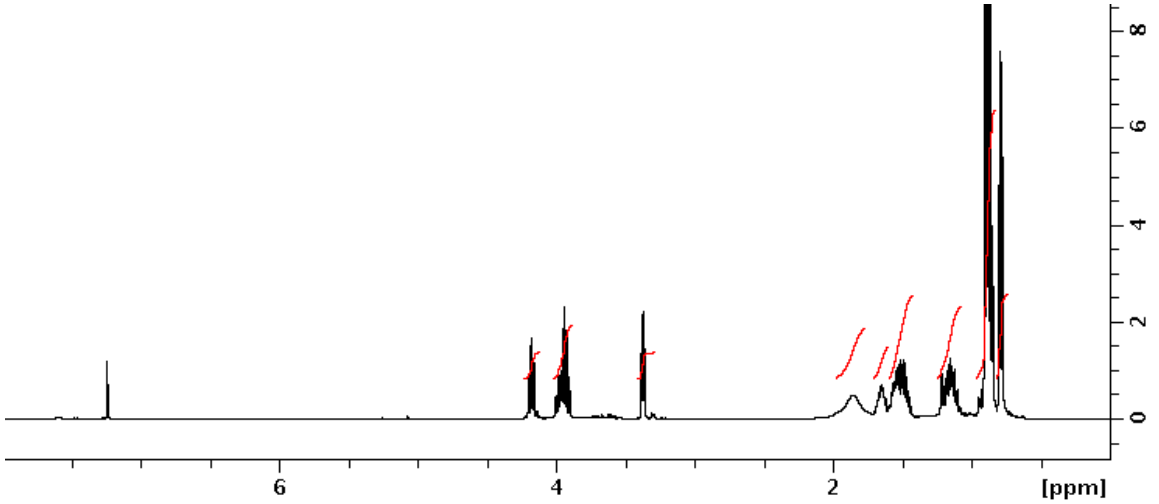


(1*S*,2*S*)-1-((*S*)-4-((*R*)-*sec*-Butyl)-4,5-dihydrooxazol-2-yl)-2-methylbutan-1-amine (L,L-**33ii**)

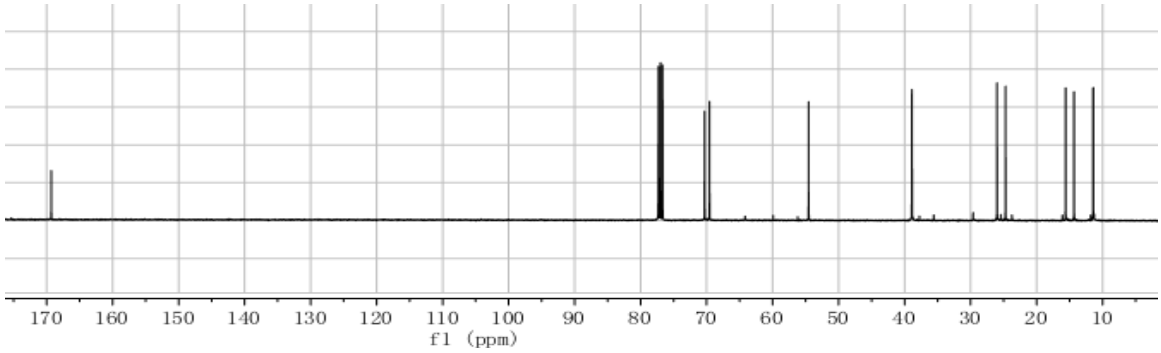


Data for L,L-33ii. Colorless oil, 390 mg (81%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 4.17 (dt, *J* = 16.0, 7.9 Hz, 1H), 3.96 (m, 2H), 3.36 (d, *J* = 7.3 Hz, 1H), 1.94 (br, 2H), 1.65 (m, 1H), 1.60 – 1.41 (m, 3H), 1.25 – 1.05 (m, 2H), 0.96 – 0.82 (m, 10H), 0.81 – 0.73 (m, 2H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 169.3, 70.3, 69.6, 54.6, 38.9, 26.0, 24.7, 15.6, 14.3, 11.5, 11.4. HRMS (ESI) m/z calcd for $C_{12}H_{25}N_2O^+$ 213.1961; found 213.1971 $(M+H)^+$.

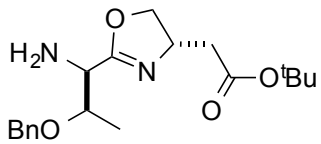
^{1}H NMR spectrum:



^{13}C NMR spectrum:



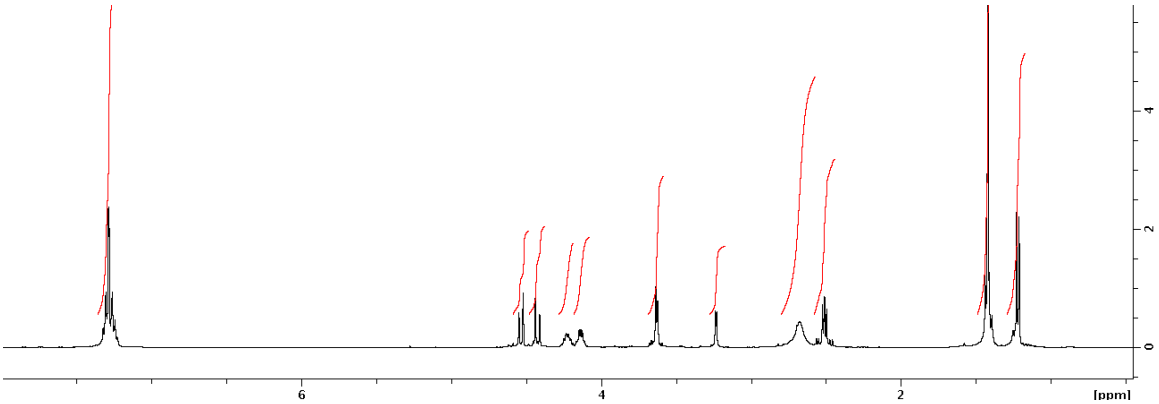
tert-Butyl 2-((*S*)-2-((1*R*,2*R*)-1-amino-2-(benzyloxy)propyl)-4,5-dihydrooxazol-4-yl)acetate (D,L-**33t'd'**)



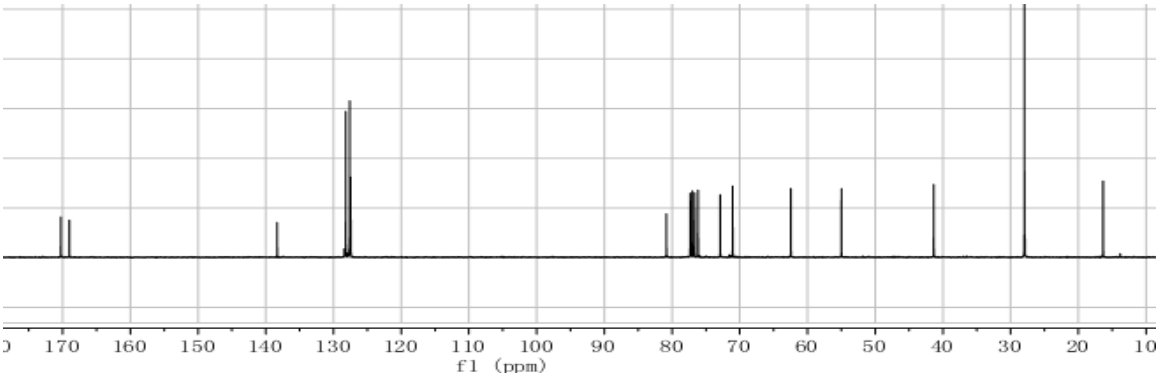
Data for D,L-33t'd'. Light yellow oil, 180 mg (80%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.35 – 7.20 (m, 5H), 4.52-4.43 (dd, *J* = 11.5, 6.1 Hz, 2H), 4.29 – 4.18 (m, 1H), 4.18 – 4.09 (m, 1H), 3.69 – 3.58 (m, 2H), 3.27 – 3.20 (m, 1H), 2.71 (br, 3H), 2.57 – 2.44 (m, 1H), 1.45 – 1.38 (m, 9H), 1.26 – 1.18 (m, 3H). ^{13}C NMR (100 MHz, $CDCl_3$) δ 173.3, 170.7, 138.3,

128.3, 127.7, 127.6, 81.1, 76.2, 75.1, 72.7, 71.3, 64.5, 55.0, 41.7, 28.0, 16.7. HRMS (ESI) m/z calcd for $C_{19}H_{29}N_2O_4^+$ 349.2122; found 349.2156 $(M+H)^+$.

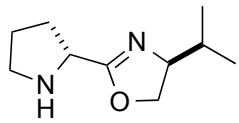
^{1}H NMR spectrum:



^{13}C NMR spectrum:



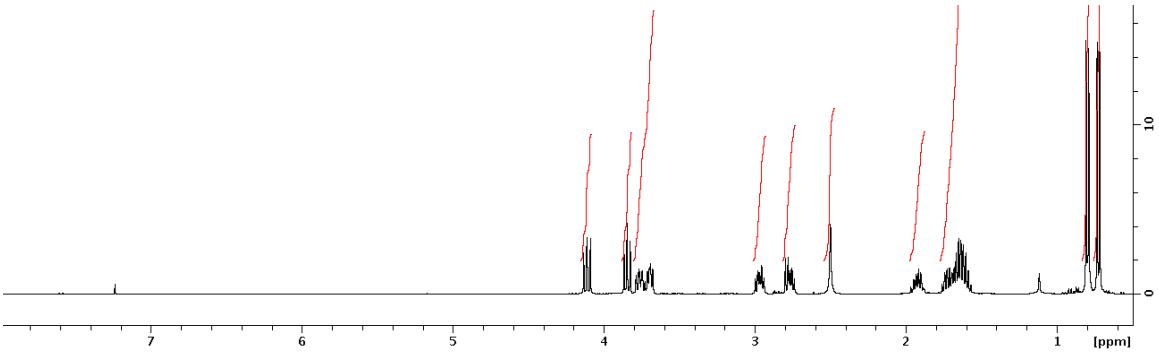
(*S*)-4-*iso*-Propyl-2-((*R*)-pyrrolidin-2-yl)-4,5-dihydrooxazole (D,L-**33pv**)



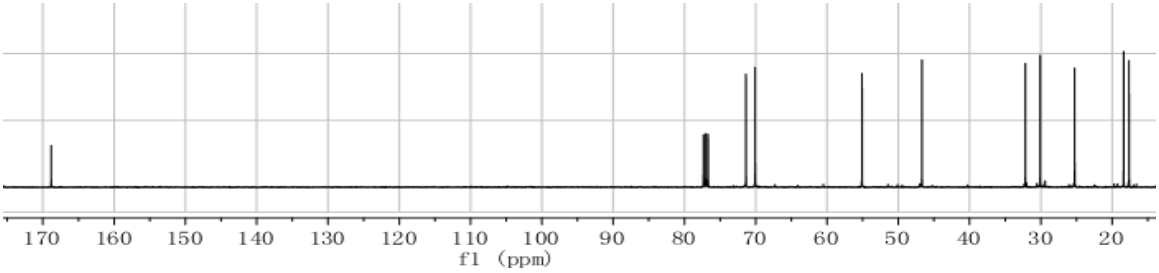
Data for D,L-33pv. Light yellow oil, 190 mg (68%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 4.11 (m, 1H), 3.85 (t, *J* = 8.0 Hz, 1H), 3.80 – 3.74 (m, 1H), 3.74 – 3.67 (m, 1H), 2.97 (m, 1H), 2.81 – 2.73 (m, 1H), 2.50 (br, 1H), 1.98 – 1.88 (m, 1H), 1.77 – 1.68 (m, 1H), 1.68 – 1.56 (m, 2H), 0.82 – 0.78 (d, *J* = 6.8 Hz, 3H), 0.73 (d, *J* = 6.8 Hz, 3H). ^{13}C NMR (100 MHz,

$CDCl_3$) δ 168.8, 71.4, 70.1, 55.1, 46.7, 32.2, 30.1, 25.3, 18.4, 17.6. MS (ESI) m/z calcd for $C_{10}H_{19}N_2O^+$ 183.1492; found 183.1786 $(M+H)^+$

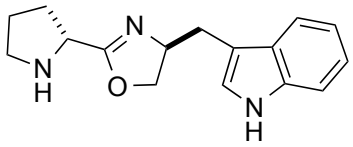
^1H NMR spectrum:



^{13}C NMR spectrum:

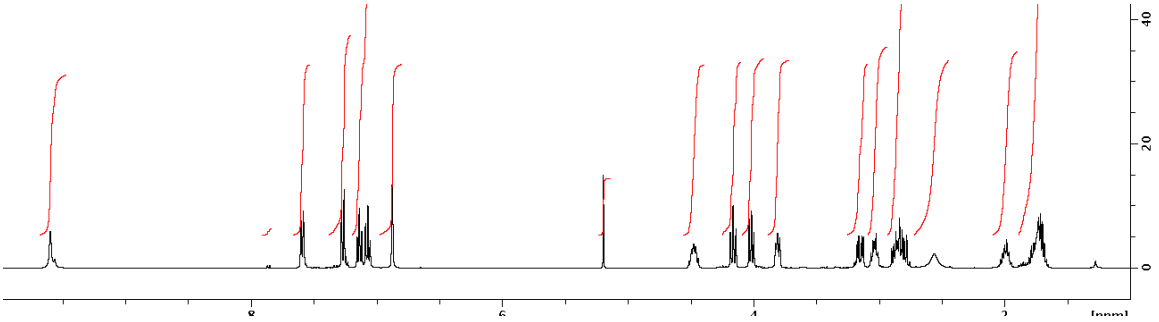


3-(((*S*)-2-((*R*)-Pyrrolidin-2-yl)-4,5-dihydrooxazol-4-yl)methyl)-1*H*-indole (D,L-**33pw**)

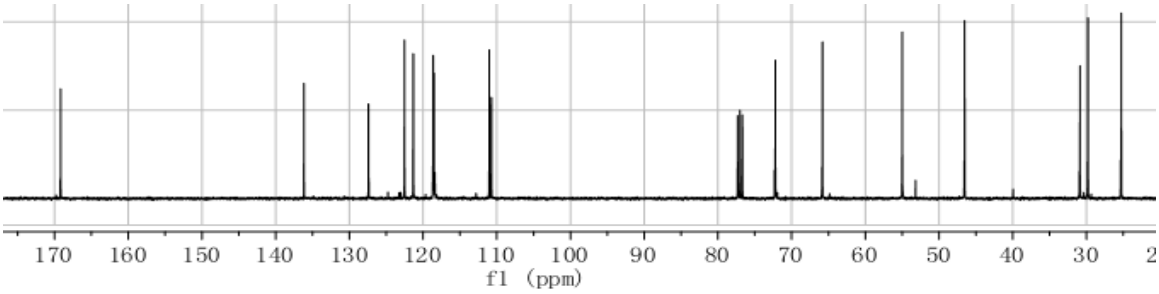


Data for D,L-33pw. Yellow oil, 380 mg (59%). ^1H NMR (400 MHz, $CDCl_3$) δ 9.58 (br, 1H), 7.58 (d, *J* = 8.0 Hz, 1H), 7.25 (d, *J* = 6.6 Hz, 1H), 7.13 - 7.07 (m, 2H), 6.87 (s, 1H), 4.47 (m, 1H), 4.16 (t, *J* = 8.9 Hz, 1H), 4.05 – 3.98 (m, 1H), 3.81 (m, 1H), 3.16 (m, 1H), 3.08 – 2.99 (m, 1H), 2.91 – 2.75 (m, 2H), 2.56 (br, 1H), 2.05 – 1.93 (m, 1H), 1.88 – 1.64 (m, 3H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 169.2, 136.2, 127.4, 122.5, 121.3, 118.6, 118.5, 118.4, 110.7, 72.2, 65.8, 55.0, 46.5, 30.9, 29.8, 25.3. HRMS (ESI) m/z calcd for $C_{16}H_{20}N_3O^+$ 270.1601; found 270.1614 $(M+H)^+$.

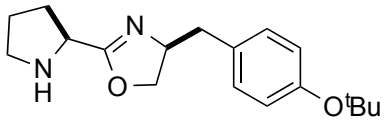
^{1}H NMR spectrum:



^{13}C NMR spectrum:

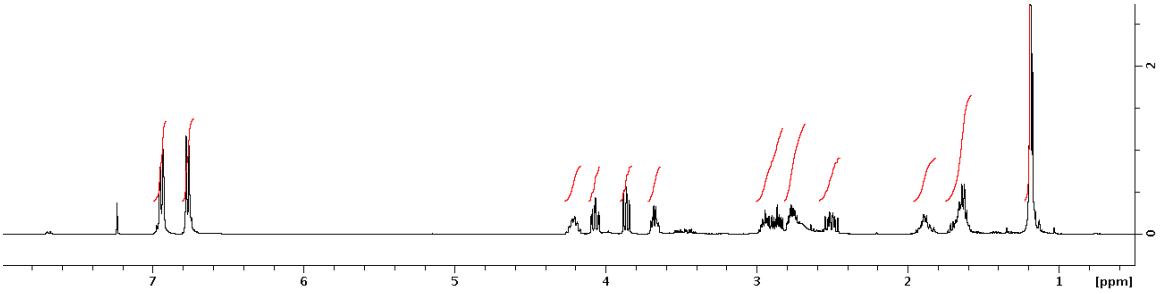


(*S*)-4-(4-*tert*-Butoxybenzyl)-2-((*S*)-pyrrolidin-2-yl)-4,5-dihydrooxazole (L,L-**33py'**)

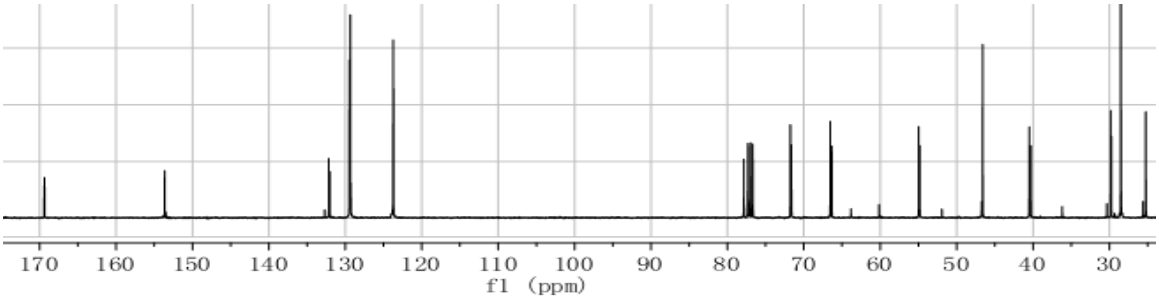


Data for L,L-33py'. Light yellow oil, 1.45 g (70%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 6.94 (dd, *J* = 8.5, 2.0 Hz, 2H), 6.76 (d, *J* = 7.1 Hz, 2H), 4.26 – 4.16 (m, 1H), 4.07 (m, 1H), 3.87 (dd, *J* = 8.3, 7.3 Hz, 1H), 3.72 – 3.63 (m, 1H), 2.92 (m, 2H), 2.81 – 2.63 (m, 2H), 2.57 – 2.45 (m, 1H), 1.96 – 1.81 (m, 1H), 1.75 – 1.57 (m, 3H), 1.18 (d, *J* = 5.7 Hz, 9H). ^{13}C NMR (100 MHz, $CDCl_3$) δ 169.4, 153.6, 132.5, 129.5, 129.4, 123.7, 77.9, 71.8, 66.5, 55.0, 46.6, 40.5, 29.8, 28.5, 25.3. HRMS (ESI) m/z calcd for $C_{18}H_{27}N_2O_2^+$ 303.2067; found 303.2062 $(M+H)^+$.

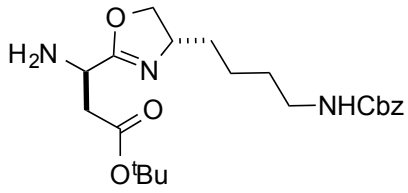
^{1}H NMR spectrum:



^{13}C NMR spectrum:

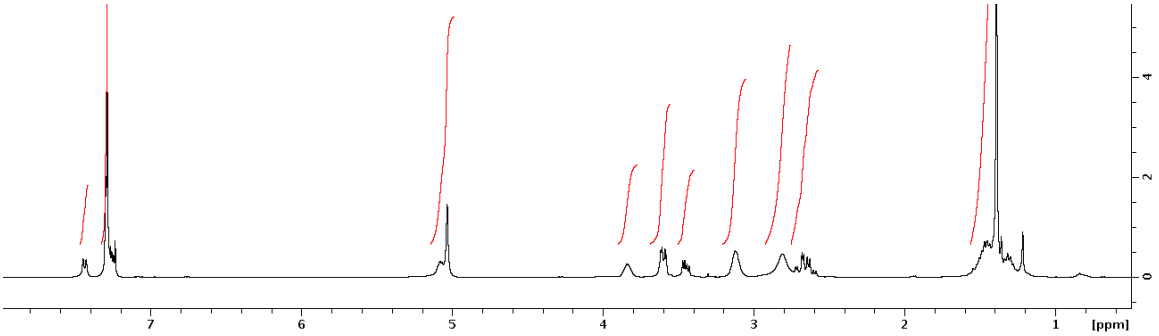


(*R*)-*tert*-Butyl 3-amino-3-((*S*)-4-(4-(benzyloxycarbonyl)butyl)-4,5-dihydrooxazol-2-yl)propanoate (D,L-**33d'k'**)

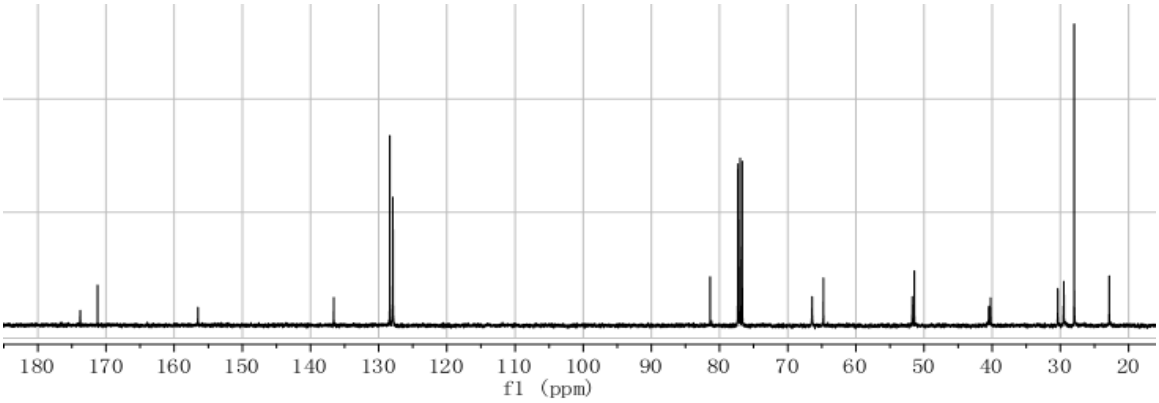


Data for D,L-33d'k'. Light yellow oil, 1.55 g (77%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.44 (d, *J* = 8.4 Hz, 1H), 7.34 – 7.25 (m, 5H), 5.13 – 5.05 (m, 2H), 3.84 (m, 1H), 3.60 (dd, *J* = 11.2, 3.4 Hz, 2H), 3.45 (dd, *J* = 11.3, 5.8 Hz, 1H), 3.12 (br, 2H), 2.81 (br, 2H), 2.74 – 2.58 (m, 2H), 1.58 –1.24 (m, 15H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 173.8, 171.2, 156.5, 136.6, 128.4, 127.9, 127.5, 81.4, 66.4, 64.8, 51.7, 51.4, 40.5, 40.2, 30.4, 29.5, 28.0, 22.8; MS (ESI) m/z calcd for $C_{22}H_{34}N_3O_5^+$ 420.2493; found 420.2800 $(M+H)^+$.

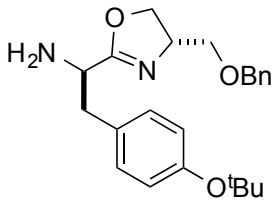
^{1}H NMR spectrum:



^{13}C NMR spectrum:



(*R*)-1-((*S*)-4-(Benzyloxymethyl)-4,5-dihydrooxazol-2-yl)-2-(4-*tert*-butoxyphenyl)ethanamine (D,L-**33y's'**)

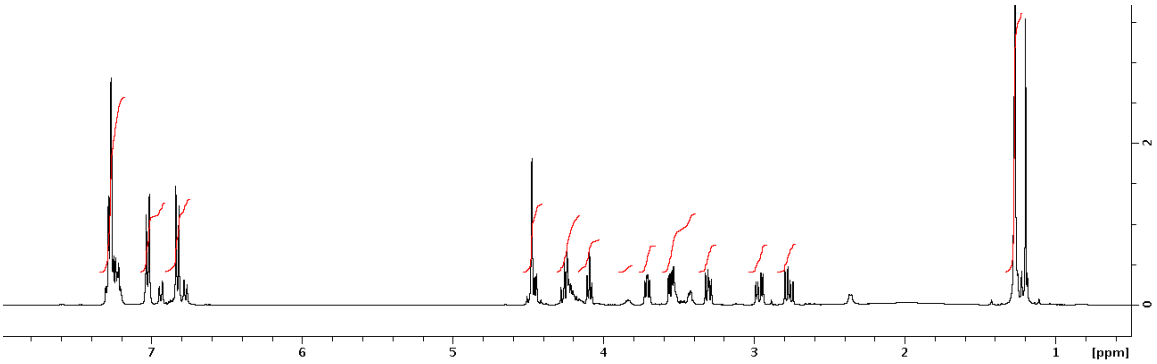


Data for D,L-33y's'. Light yellow oil, 292 mg (62%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.32 – 7.19 (m, 5H), 7.03 (d, *J* = 8.4 Hz, 2H), 6.83 (d, *J* = 8.4 Hz, 2H), 4.50 – 4.43 (m, 2H), 4.29 – 4.18 (m, 3H), 4.09 (t, *J* = 6.5 Hz, 1H), 3.71 (dd, *J* = 7.5, 5.5 Hz, 1H), 3.58 – 3.39 (m, 2H), 3.31 (dd, *J* = 9.3, 6.6 Hz, 1H), 2.97 (dd, *J* = 13.6, 5.2 Hz, 1H), 2.77 (dd, *J* = 13.6,

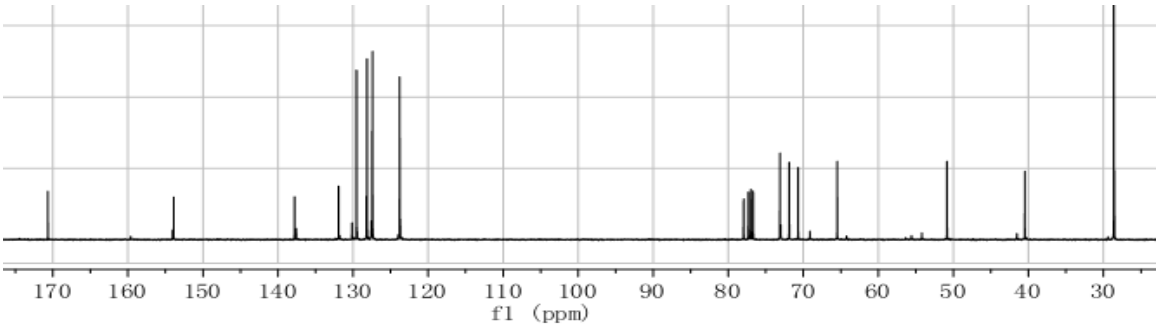
7.8 Hz, 1H), 1.26 (d, *J* = 7.4 Hz, 9H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 170.7, 153.9, 137.8, 131.9, 128.7, 128.2, 127.5, 123.8, 123.7, 77.9, 73.1, 71.9, 70.7, 65.5, 50.9, 40.5, 28.6. MS (ESI) m/z calcd for $C_{23}H_{30}N_2O_3Li^+$ 389.2411; found 389.1712 $(M+Li)^+$.

\

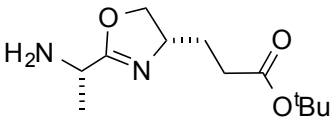
^{1}H NMR spectrum:



^{13}C NMR spectrum:

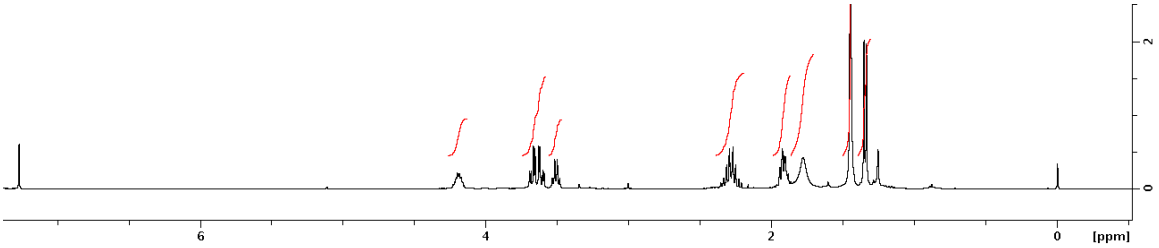


tert-Butyl 3-((*S*)-2-((*S*)-1-aminoethyl)-4,5-dihydrooxazol-4-yl)propanoate (L,L-**33ae'**)

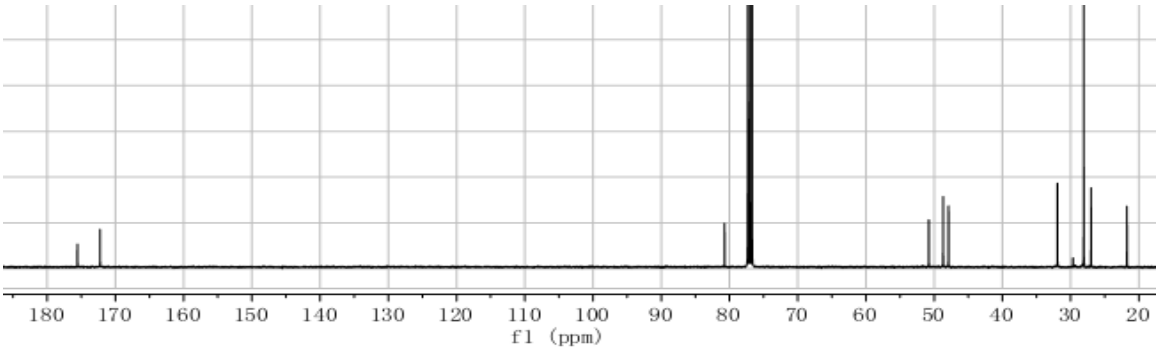


Data for L,L-33ae'. Yellow oil, 1.07 g (72%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 4.26 – 4.12 (m, 1H), 3.64 (ddd, *J* = 25.4, 11.2, 3.9 Hz, 2H), 3.51 (q, *J* = 6.9 Hz, 1H), 2.38 – 2.19 (m, 2H), 1.98 – 1.85 (m, 2H), 1.81 (br, 2H), 1.45 (s, 9H), 1.35 (d, *J* = 8.2 Hz, 3H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 175.6, 172.3, 80.7, 50.8, 48.7, 47.9, 31.9, 28.1, 27.0, 21.8. MS (ESI) m/z calcd for $C_{12}H_{23}N_2O_3^+$ 243.1703; found 243.1986 $(M+H)^+$.

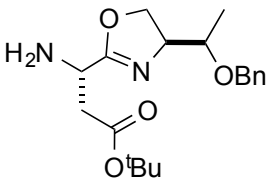
^{1}H NMR spectrum:



^{13}C NMR spectrum:

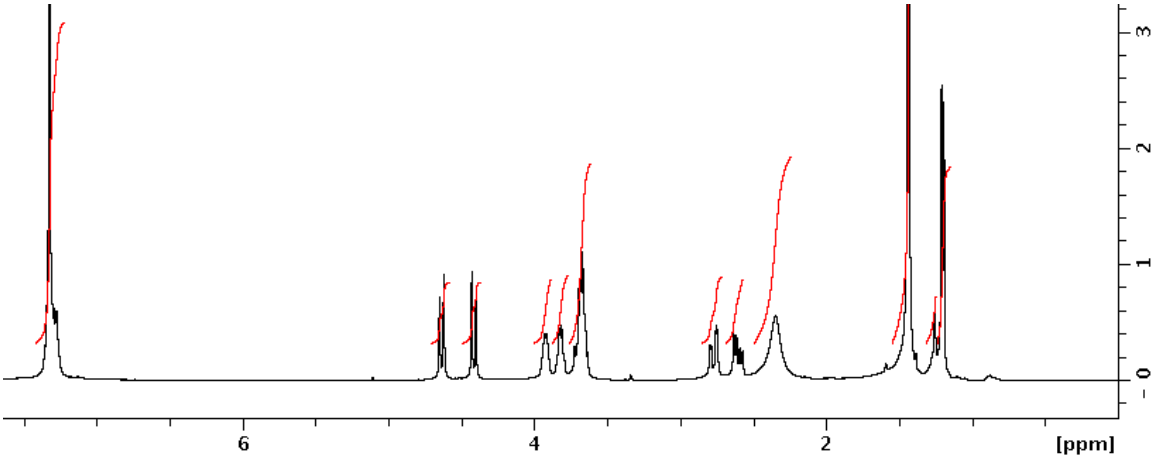


(*S*)-*tert*-Butyl 3-amino-3-((*S*)-4-((*S*)-1-(benzyloxy)ethyl)-4,5-dihydrooxazol-2-yl)propanoate (L,D-**33d't'**)

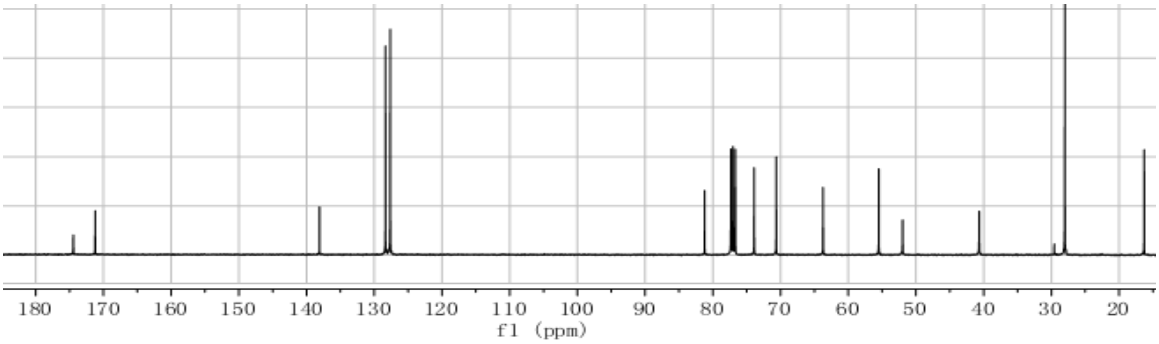


Data for L,D-33d't'. Yellow oil, 120 mg (60%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.37 – 7.24 (m, 5H), 4.63 (d, *J* = 11.7 Hz, 1H), 4.41 (d, *J* = 11.7 Hz, 1H), 3.97 – 3.88 (m, 1H), 3.82 (m, 1H), 3.68 (m, 3H), 2.77 (dd, *J* = 16.6, 3.6 Hz, 1H), 2.60 (dd, *J* = 16.6, 7.5 Hz, 1H), 2.35 (br, 2H), 1.43 (s, 9H), 1.20 (d, *J* = 6.2 Hz, 3H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 174.4, 171.2, 138.1, 128.4, 127.7, 127.6, 81.2, 73.9, 70.6, 63.7, 55.5, 52.0, 40.7, 28.0, 16.3. HRMS (ESI) m/z calcd for $C_{19}H_{29}N_2O_4^+$ 349.2122; found 349.2127 $(M+H)^+$.

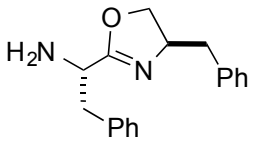
^{1}H NMR spectrum:



^{13}C NMR spectrum:



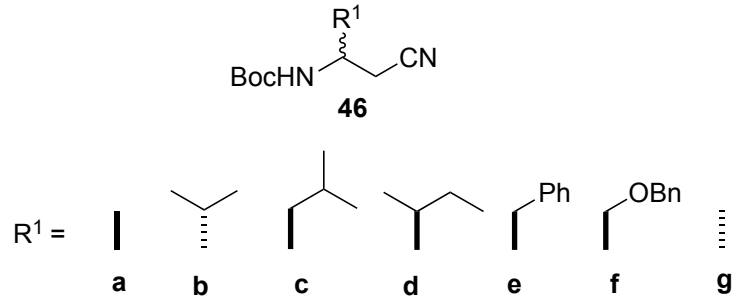
Procedures for the preparation of (S)-1-((S)-4-benzyl-4,5-dihydrooxazol-2-yl)-2-phenylethanamine (L,D-**33ff**)



To a solution of Cbz-containing oxazoline compound L,D-**32ff** (0.45 g, 1.09 mmol) in anhydrous methanol (10 mL) under nitrogen was added 10 wt % Pd/C (0.1 equiv Pd, 109 mg). The reaction was placed under an atmosphere of hydrogen (1 atm, balloon) for 12 h. After the completion of reaction (as monitored by TLC), the flask was purged with

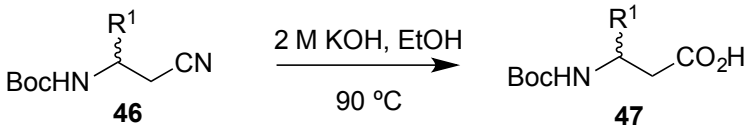
N_2. The reaction mixture was filtered over a celite pad and washed with portions of methanol (3 × 5 mL), the filtrate and washings were collected and concentrated under vacuum to afford the crude amine product L,D-**33ff** as colorless oil (570 mg, 90 %), which can be directly used for further reaction without any purification.

*General Procedure for the Preparation of N-Boc Amino Nitrile (**46**)*



N-Boc-protected amino alcohol was prepared from *N*-Boc-protected amino acid using known literature procedures.[198] The crude *N*-protected amino nitrile compounds were synthesized on large scale from corresponding *N*-Boc-protected amino alcohols using the procedures from the literature.[199]

*General Procedure for the Preparation of N-Boc-β-Amino Acid (**47**)*



Typically, *N*-protected amino nitrile compounds **46** (20 mmol) was dissolved in EtOH (20 mL) at room temperature and 2 M KOH aqueous solution (20 mL) was added in one

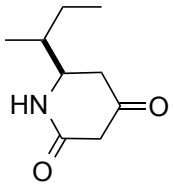
portion. The resulting mixture was stirred at 90 °C until the reaction was complete, as determined by TLC (usually 3 h was enough). The resulting clear mixture was allowed to cool down to room temperature and EtOH was removed under vacuum. The crude product was cooled to 0 °C and 1M HCl aqueous solution was slowly added until the pH of the solution reached about 1 - 2. The solution was extracted with Et_2O three times (3 × 80 mL). The combined Et_2O solution was washed twice with water (2 × 100 mL) and brine (100 mL). The ether phase was dried over anhydrous $MgSO_4$ and concentrated under vacuum. The resulting oily residue was used directly in the synthesis of compounds **34** without any purification.

*General Procedure for the Preparation of Piperidine-2,4-dione (**34**)*[191,192,200]

To a stirred solution of meldrum's acid (476 mg, 3.3 mmol) and DMAP (550 mg, 4.5 mmol) at 0 ºC in dichloromethane (15 mL) was added *N*-Boc-β-amino acid compounds **47** (3.0 mmol) in one portion. EDCI (978 mg, 5.1 mmol) was added in one portion and the reaction mixture was stirred at 25 ºC for 3 - 5 h. The yellow reaction mixture was transferred to a separatory funnel, diluted with dichloromethane (80 mL) and washed with cold 5 % $KHSO_4$ (3 × 75 mL) and brine (75 mL). The organic layer was dried over $MgSO_4$ and filtered. The filtrate was concentrated under vacuum and 60 mL ethyl acetate was added. After refluxing for 5 h under N_2 atmosphere, the solution was dissolved in 1:1 TFA/dichloromethane to give 0.1 M concentration at 0 °C, and the reaction was stirred for 30 min at room temperature. Toluene (30 mL) was added and the solution was concentrated. Residual TFA was azeotroped 3 times with toluene (3 × 30 mL) and the residue was placed under high vacuum for 3 h to give the crude mixture. The crude product was dissolved in CH_2Cl_2 and a small portion of anhydrous K_2CO_3 was added. The solid was filtered and the filtrate was concentrated under reduced pressure.

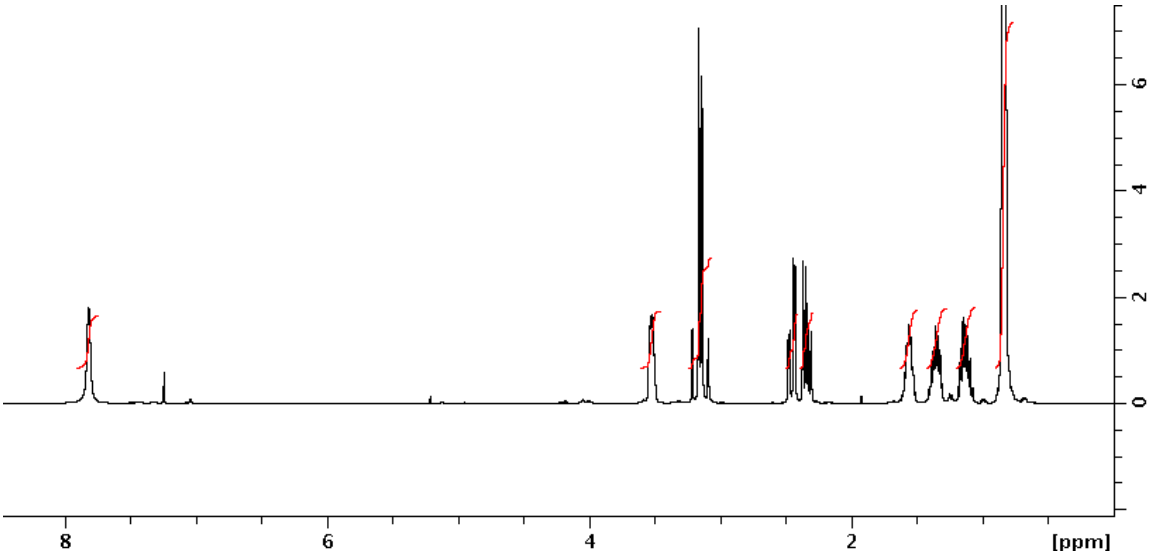
The crude product was purified by flash chromatography (CH_2Cl_2–MeOH 98:2 to 95:5) to provide the desired product.[55]

(*R*)-6-((*S*)-*sec*-Butyl)piperidine-2,4-dione (L-**34i**)

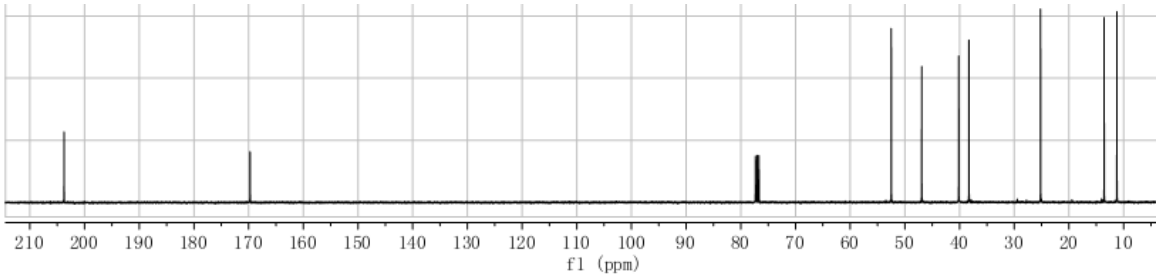


Data for L-34i. Pale yellow crystal, 60 mg (54%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.82 (br, 1H), 3.52 (m, 1H), 3.23 – 3.07 (m, 2H), 2.40 (dd, *J* = 16.1, 6.7 Hz, 1H), 2.30 (dd, *J* = 16.1, 6.7 Hz, 1H), 1.62 – 1.50 (m, 1H), 1.38 (m, 2H), 1.19 (m, 2H), 0.89 – 0.79 (m, 6H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 203.7, 169.7, 52.5, 46.9, 40.2, 38.3, 25.2, 13.6, 11.2; HRMS (ESI) m/z calcd for $C_9H_{16}NO_2^+$ 170.1176; found 170.1177 (M+H)+. IR (film, cm^{-1}) 2924, 2859, 2361, 1728, 1667.

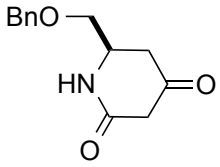
^{1}H NMR spectrum:



^{13}C NMR spectrum:

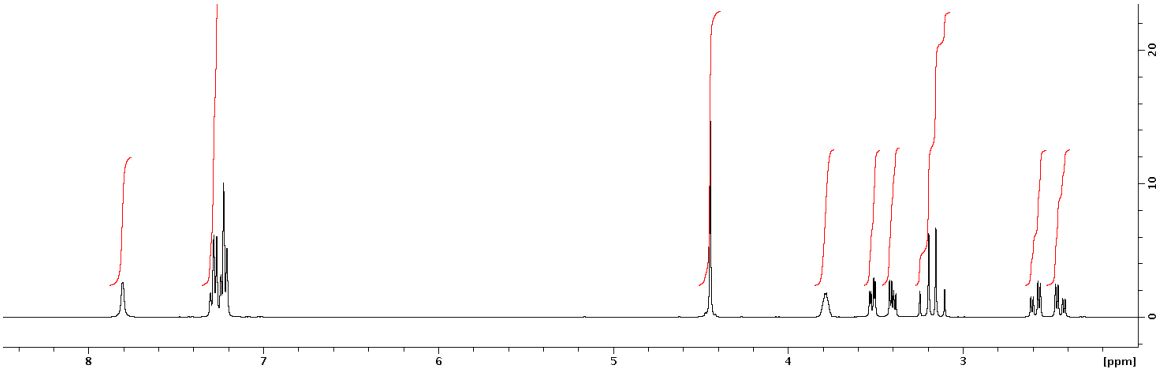


(*R*)-6-(Benzyloxymethyl)piperidine-2,4-dione (L-**34s**)

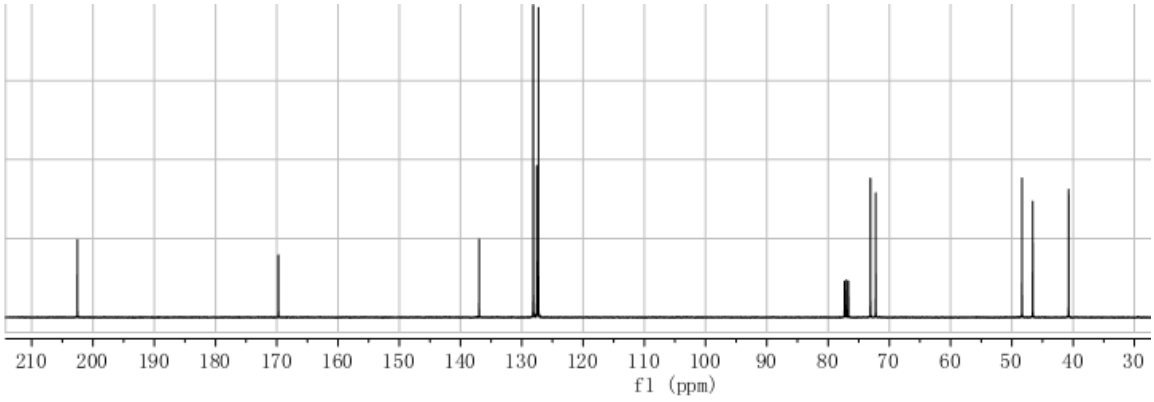


Data for L-34s. Pale yellow crystal, 1.05 g (65%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.81 (s, 1H), 7.26 (m, 5H), 4.45 (s, 2H), 3.82 – 3.75 (m, 1H), 3.55 – 3.37 (m, 2H), 3.22 – 3.15 (dd, *J* = 36.4, 20.4 Hz, 2H), 2.57 (dd, *J* = 16.2, 5.8 Hz, 1H), 2.45 (dd, *J* = 16.2, 5.8 Hz, 1H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 202.6, 169.7, 137.0, 128.1, 127.6, 127.3, 73.1, 72.2, 48.4, 46.6, 40.7; MS (ESI) m/z calcd for $C_{13}H_{16}NO_3^+$ 234.1125; found 234.1138 $(M+H)^+$; IR (film, cm^{-1}) 3233, 2924, 2859, 1724, 1670.

^{1}H NMR spectrum:



^{13}C NMR spectrum:



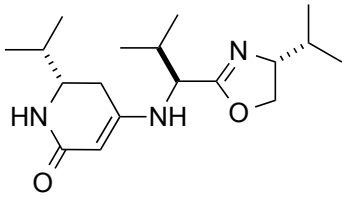
*General Procedure for the Preparation of Scaffold **30**.*

	R^1	R^2	R^3	Yield%
D,L,D-**30vvv**	D-Val	L-Val	D-Val	63 (Method A)
L,D,L-**30fpv**	L-Phe	D-Pro	L-Val	40 (Method A)
L,D,L-**30lpw**	L-Leu	D-Pro	L-Trp	35 (Method B)
L,L,L-**30ipy'**	L-Ile	L-Pro	L-Tyr(^{t}Bu)	31 (Method A) 65 (Method B)
L,D,L-**30s'y's'**	L-Ser(Bn)	D-Tyr(^{t}Bu)	L-Ser(Bn)	38 (Method A)
L,L,L-**30fii**	L-Phe	L-Ile	L-Ile	88 (Method B)

A representative synthesis of the D,L,D-**30vvv** is described (method A) using a modified reported procedure.[201,202] To a solution of compound L,D-**33vv** (0.110 g, 0.597 mmol) in dry isopropyl alcohol (4.5 mL) was added compound D-**34v** (0.102 g, 0.657 mmol), followed by acetic acid (0.1 mL). The resulting solution was allowed to react at 55 ºC for 15 min, and dried 3Å molecular sieve was added to the solution (the 3Å MS should be put in a 300 ºC muffle furnace for 24 h before using). The reaction was stirred at 55 ºC

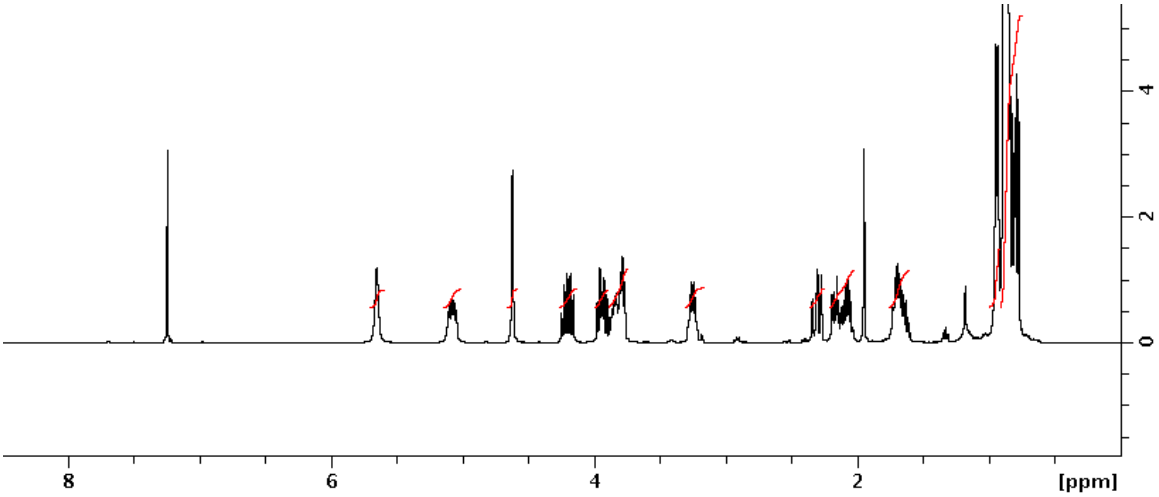
for another 45 min until completion, as monitored by TLC. The solvent was removed under vacuum and the oily residue was purified by flash chromatography (silica gel, CH_2Cl_2-MeOH, gradually from 98:2 to 92:8) to afford the peptidomimetic compound D,L,D-**30vvv** as pale yellow oil (120 mg, 63 %).

(*S*)-6-*iso*-Propyl-4-((*S*)-1-((*R*)-4-*iso*-propyl-4,5-dihydrooxazol-2-yl)-2-methylpropylamino)-5,6-dihydropyridin-2(1*H*)-one (D,L,D-**30vvv**)

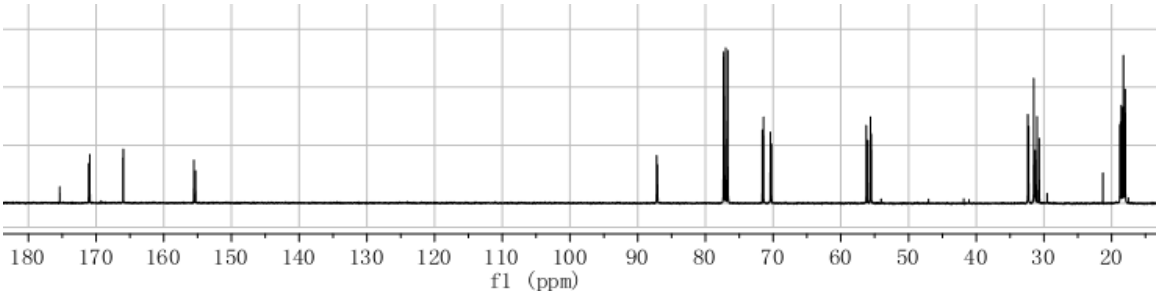


Data for D,L,D-30vvv: ^{1}H NMR (400 MHz, $CDCl_3$) δ 5.65 (s, 1H), 5.07 (m, 1H), 4.61 (s, 1H), 4.26 – 4.14 (m, 1H), 3.94 (m, 1H), 3.89 – 3.74 (m, 2H), 3.31 – 3.19 (m, 2H), 2.29 (m, 1H), 2.22 – 2.14 (m, 1H), 2.15 – 2.02 (m, 1H), 1.76 – 1.59 (m, 1H), 0.99 – 0.75 (m, 18H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 175.4, 171.1, 166.0, 155.6, 86.7, 71.6, 70.4, 56.2, 55.6, 32.4, 31.3, 31.0, 18.8, 18.6, 18.2, 18.0; HRMS (ESI) m/z calcd for $C_{18}H_{32}N_3O_2^+$ 322.2489; found 322.2300 $(M+H)^+$. IR (film, cm^{-1}) 3271, 2963, 2932, 2874, 2361, 2338, 1728, 1663, 1624.

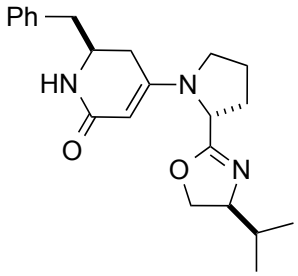
^{1}H NMR spectrum:



^{13}C NMR spectrum:



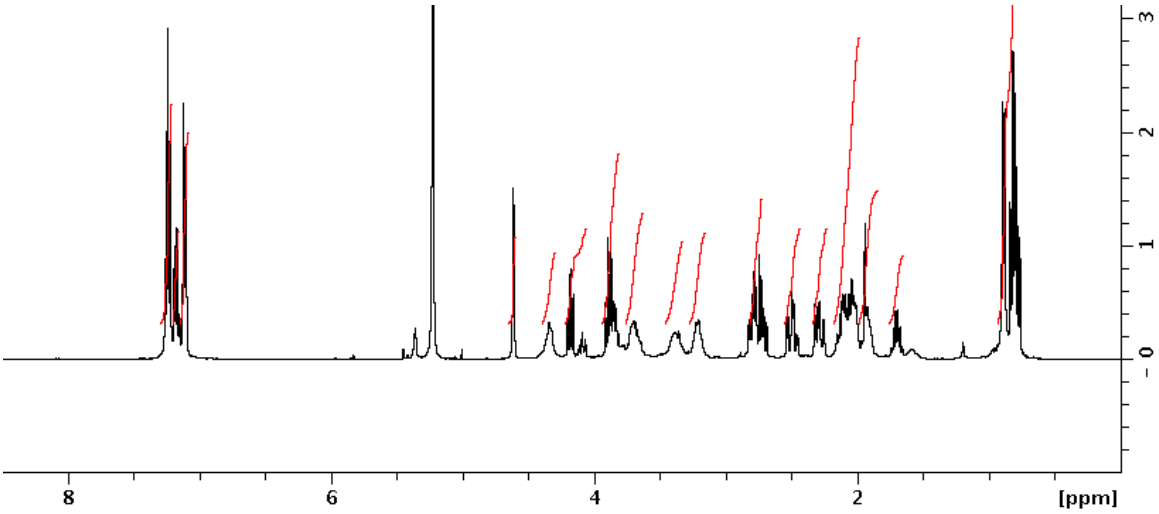
(*S*)-6-Benzyl-4-((*R*)-2-((*S*)-4-*iso*-propyl-4,5-dihydrooxazol-2-yl)pyrrolidin-1-yl)-5,6-dihydropyridin-2(1*H*)-one (L,D,L-**30fpv**)



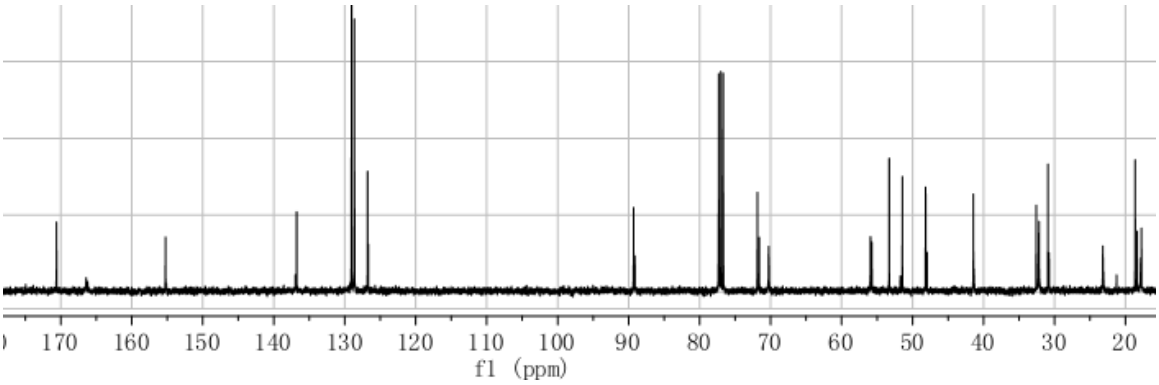
Data for L,D,L-30fpv: Brown oil, 110 mg (40%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.24- 7.11 (m, 5H), 4.61 (s, 1H), 4.34 (m, 1H), 4.22 – 4.06 (m, 1H), 3.93 – 3.61 (m, 3H), 3.37 (m,

1H), 3.22 (m, 1H), 2.84 – 2.76 (m, 1H), 2.72 (m, 1H), 2.56 – 2.44 (m, 1H), 2.29 (m, 1H), 2.17 – 1.98 (m, 3H), 1.98 – 1.87 (m, 2H), 1.75 – 1.64 (m, 1H), 0.93 – 0.74 (m, 6H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 170.6, 166.4, 155.3, 136.8, 129.0, 128.6, 126.8, 126.7, 89.3, 71.9, 70.3, 55.8, 53.3, 51.1, 48.2, 41.5, 32.2, 32.0, 30.8, 23.2, 18.7, 17.8; HRMS (ESI) m/z calcd for $C_{22}H_{30}N_3O_2^+$ 368.2333; found 368.2334 $(M+H)^+$. IR (film, cm^{-1}) 3028, 3260, 2959, 2874, 1667, 1632, 1574.

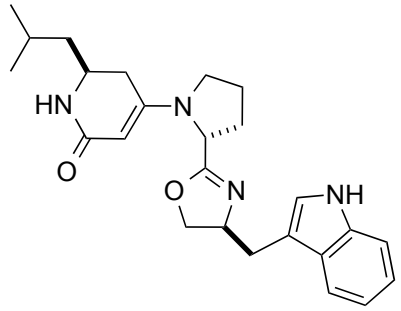
^{1}H NMR spectrum:



^{13}C NMR spectrum:



(*S*)-4-((*R*)-2-((*S*)-4-((1*H*-Indol-3-yl)methyl)-4,5-dihydrooxazol-2-yl)pyrrolidin-1-yl)-6-*iso*butyl-5,6-dihydropyridin-2(1*H*)-one (L,D,L-**30lpw**)

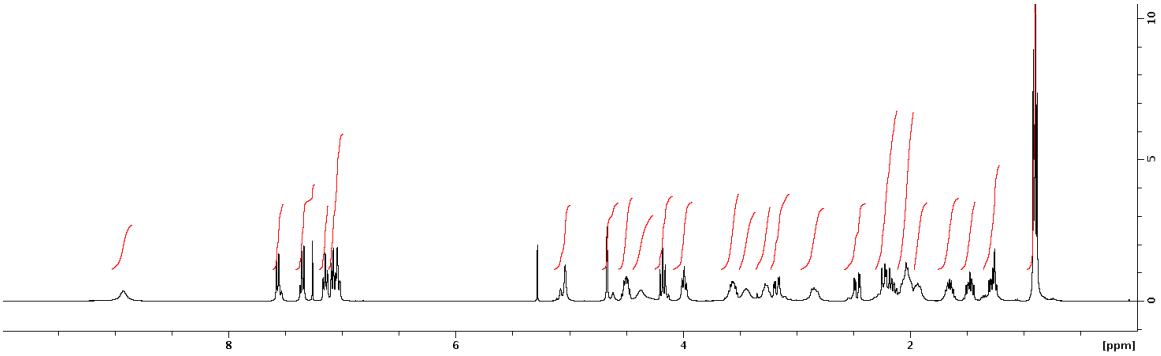


A representative synthesis of the L,D,L-**30lpw** using method B is as follows. To a solution of compound D,L-**33pw** (25 mg, 0.093 mmol) in dry dicholoromethane (5 mL), compound L-**34l** (18.85 mg, 0.111 mmol) was added, followed by $KHCO_3$ (93 mg, 0.928 mmol), HOAt (25.3 mg, 0.186 mmol) and EDCI (35.6 mg, 0.186 mmol). The resulting suspension was allowed to stir at 25 ºC for 14 h until completion, as monitored by TLC. The solvent was removed under vacuum and the residue was dissolved in dicholoromethane (15 mL), washed with saturated $NaHCO_3$ solution (2 × 10 mL) and brine (10 mL). The combined solvent was concentrated under vacuum and purified by flash chromatography (CH_2Cl_2-MeOH, gradually from 98:2 to 92:8) to afford the final peptidomimetic compound L,D,L-**30lpw** as yellow oil (22.5 mg, 35 %).

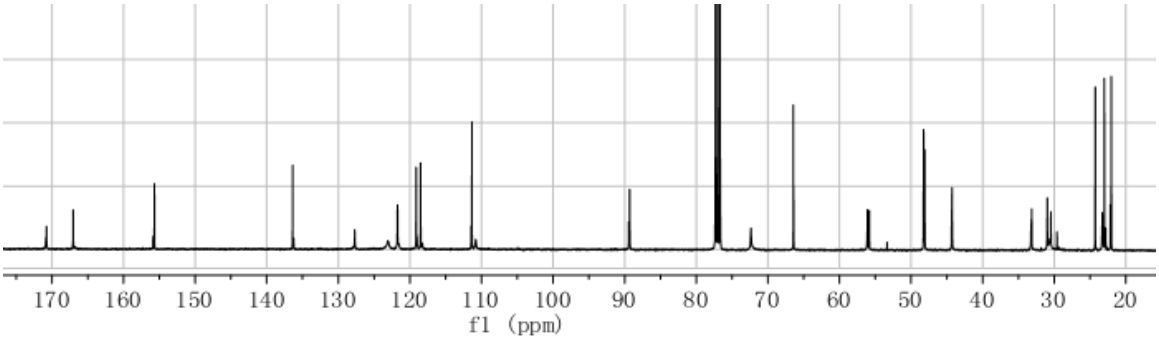
Data for compound L,D,L-30lpw: Yellow oil, ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.93 (br, 1H), 7.56 (d, *J* = 8.9 Hz, 1H), 7.39 – 7.32 (m, 1H), 7.15 (t, *J* = 7.5 Hz, 1H), 7.11 – 7.00 (m, 2H), 5.11 – 5.00 (m, 1H), 4.64 (s, 1H), 4.56 – 4.45 (m, 1H), 4.37 (br, 1H), 4.16 (dd, *J* = 19.6, 10.4 Hz, 1H), 3.99 (t, *J* = 7.9 Hz, 1H), 3.64 – 3.51 (m, 1H), 3.44 (m, 1H), 3.33 – 3.21 (m, 1H), 3.17 (dd, *J* = 14.6, 4.5 Hz, 1H), 2.84 (m, 1H), 2.47 (dd, *J* = 16.0, 4.7 Hz, 1H), 2.17 (m, 2H), 2.11 – 1.97 (m, 2H), 1.93 (m, 1H), 1.64 (m, 1H), 1.53 – 1.42 (m, 1H),

1.27 (m, 1H), 0.90 (m, 6H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 170.9, 167.0, 155.9, 136.4, 127.7, 123.1, 121.7, 119.1, 118.5, 118.3, 111.4, 111.3, 110.8, 89.4, 72.4, 66.5, 56.1, 55.8, 48.3, 44.3, 33.2, 30.6, 24.3, 23.2, 22.0; HRMS (ESI) m/z calcd for $C_{25}H_{33}N_4O_2^+$ 421.2598; found 421.2534 $(M+H)^+$. IR (film, cm^{-1}) 3213, 3051, 2955, 2924, 2870, 2361, 1663, 1612, 1574.

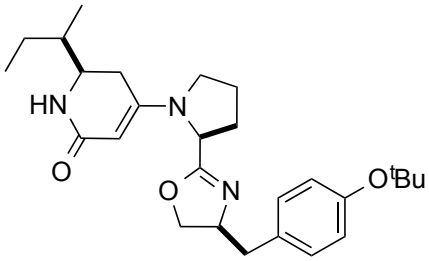
^{1}H NMR spectrum:



^{13}C NMR spectrum:

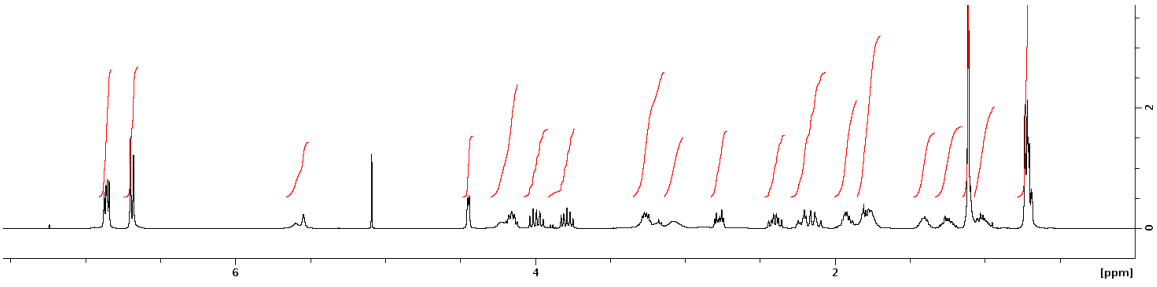


(*R*)-4-((*S*)-2-((*S*)-4-(4-*tert*-Butoxybenzyl)-4,5-dihydrooxazol-2-yl)pyrrolidin-1-yl)-6-(2-methylbutyl)-5,6-dihydropyridin-2(1*H*)-one (L,L,L-**30ipy'**)

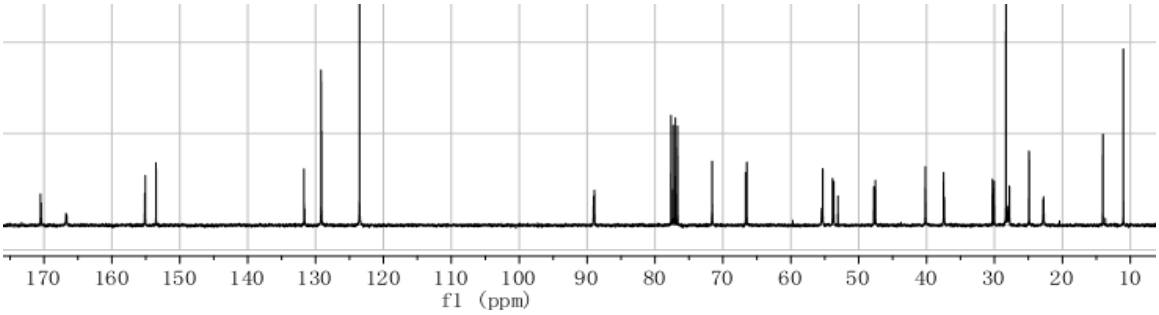


Data for L,L,L-30ipy': Isolated as dark brown oil, 34 mg (31%, method A); 540 mg (65%, method B). ^{1}H NMR (400 MHz, CDCl3) δ 6.86 (dd, *J* = 8.3, 5.0 Hz, 2H), 6.69 (d, *J* = 8.4 Hz, 2H), 5.57 (m, 1H), 4.44 (m, 1H), 4.17 (m, 2H), 4.06 – 3.92 (m, 1H), 3.85 – 3.73 (m, 1H), 3.33 – 3.15 (m, 2H), 3.10 (m, 1H), 2.77 (m, 1H), 2.46 – 2.13 (m, 3H), 1.92 (m, 1H), 1.85 – 1.70 (m, 2H), 1.41 (m, 1H), 1.33 – 1.17 (m, 1H), 1.10 (d, *J* = 6.2 Hz, 9H), 1.06 – 0.94 (m, 2H), 0.77 – 0.66 (m, 6H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 170.5, 166.8, 155.2, 153.5, 131.7, 129.2, 129.1, 123.5, 89.1, 77.6, 71.6, 66.7, 55.0, 53.9, 47.8, 40.2, 37.5, 30.3, 28.3, 27.8, 24.9, 22.8, 14.1, 11.0; HRMS (ESI) m/z calcd for $C_{27}H_{40}N_3O_3^+$ 454.3064; found 454.3069 $(M+H)^+$. IR (film, cm^{-1}) 3264, 2970, 2928, 2874, 2361, 1736, 1620, 1578, 1504.

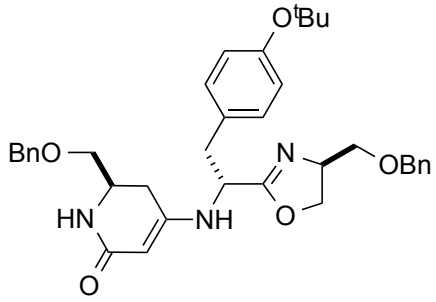
^{1}H NMR spectrum:



^{13}C NMR spectrum:

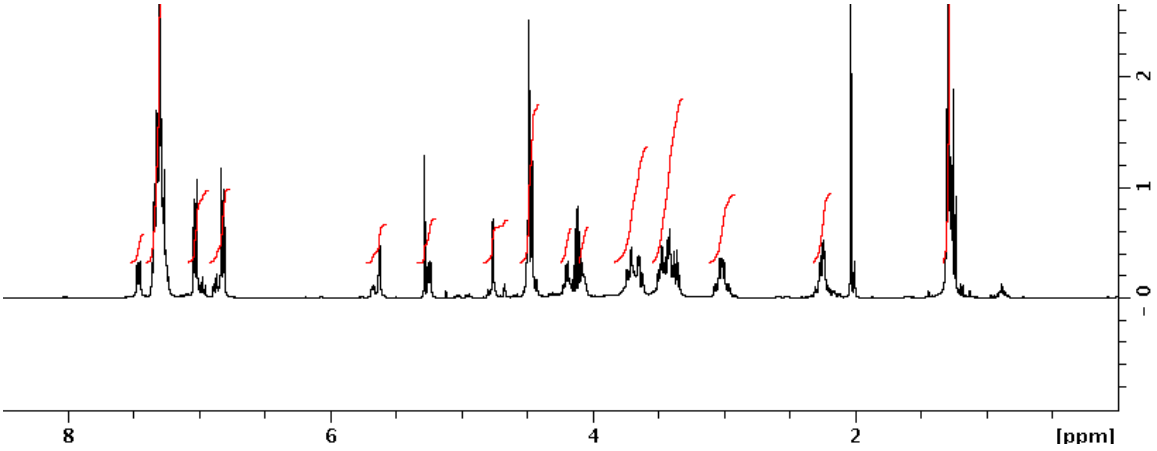


(*R*)-6-(Benzyloxymethyl)-4-((*R*)-1-((*S*)-4-(benzyloxymethyl)-4,5-dihydrooxazol-2-yl)-2-(4-tert-butoxyphenyl)ethylamino)-5,6-dihydropyridin-2(1*H*)-one (L,D,L-**30s'y's'**)

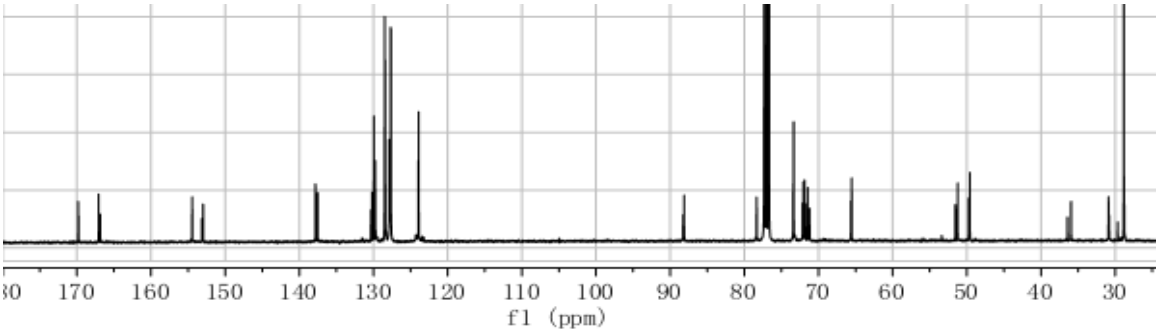


Data for L,D,L-30s'y's': Isolated as yellow oil, 22 mg (38%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.46 (d, *J* = 8.0 Hz, 1H), 7.37 – 7.22 (m, 9H), 7.06 – 6.94 (m, 2H), 6.80 – 6.78 (m, 2H), 5.70 – 5.60 (m, 1H), 5.24 (m, 1H), 4.76 (s, 1H), 4.52 – 4.41 (m, 4H), 4.20 (m, 1H), 4.11 (m, 1H), 3.79 – 3.58 (m, 3H), 3.54 – 3.31 (m, 4H), 3.09 – 2.93 (m, 2H), 2.32 – 2.18 (m, 2H), 1.34 – 1.25 (m, 9H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 169.8, 167.1, 154.5, 153.1, 137.8, 137.5, 131.6, 130.4, 130.2, 128.6, 128.4, 127.9, 127.8, 124.3, 124.0, 88.3, 78.4, 73.4, 72.1, 71.9, 71.5, 65.6, 51.6, 49.6, 36.0, 30.9, 28.8; HRMS (ESI) m/z calcd for $C_{36}H_{44}N_3O_5^+$ 598.3275; found 598.3265 $(M+H)^+$. IR (film, cm^{-1}) 3271, 3063, 2974, 2928, 2862, 2361, 1728, 1624, 1555, 1504.

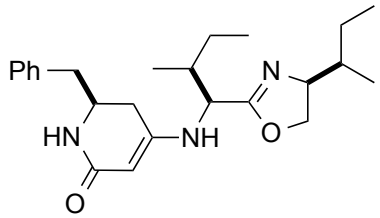
^{1}H NMR spectrum:



^{13}C NMR spectrum:



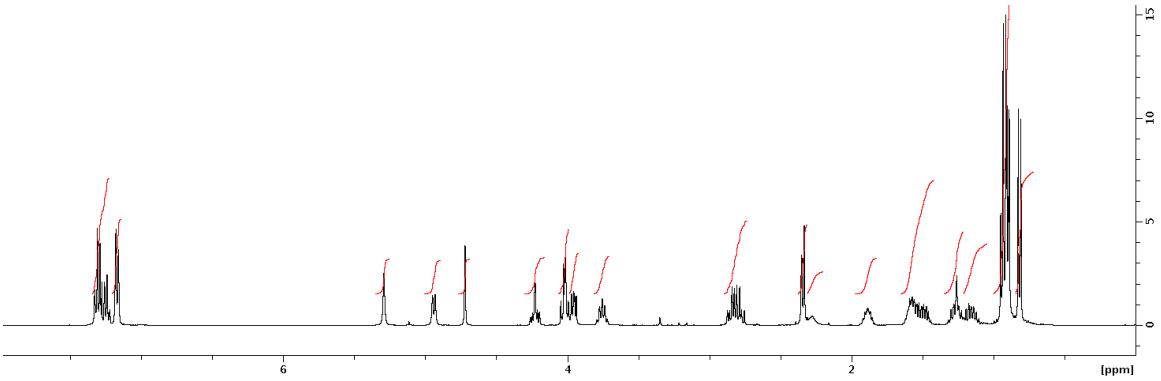
(*S*)-6-Benzyl-4-((1*S*,2*S*)-1-((*S*)-4-((*R*)-*sec*-butyl)-4,5-dihydrooxazol-2-yl)-2-methylbutylamino)-5,6-dihydropyridin-2(1*H*)-one (L,L,L-**30fii**)



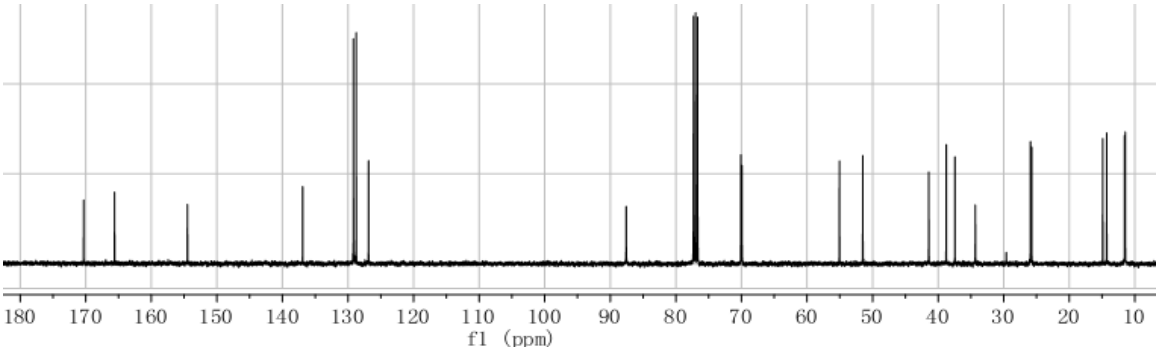
Data for L,L,L-30fii: Yellow oil, 70 mg (88%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.30 – 7.21 (m, 3H), 7.20 – 7.14 (m, 2H), 5.29 (br, 1H), 4.94 (d, *J* = 7.8 Hz, 1H), 4.72 (s, 1H), 4.27 – 4.18 (m, 1H), 4.06 – 3.99 (m, 1H), 3.95 (m, 1H), 3.81 – 3.70 (m, 1H), 2.89 – 2.74 (m, 2H), 2.34 (d, *J* = 7.5 Hz, 2H), 2.29 (m, 1H), 1.94 – 1.83 (m, 1H), 1.64 – 1.44 (m, 2H), 1.32 – 1.20 (m, 2H), 1.20 – 1.08 (m, 1H), 0.97 – 0.86 (m, 9H), 0.82 (d, *J* = 6.7 Hz, 3H);

^{13}C NMR (100 MHz, $CDCl_3$) δ 170.3, 165.6, 154.5, 136.9, 129.2, 128.8, 126.9, 87.6, 70.1, 69.9, 55.1, 51.5, 41.4, 38.8, 37.4, 34.3, 25.9, 25.8, 14.9, 14.3, 11.6, 11.5; HRMS (ESI) m/z calcd for $C_{24}H_{36}N_3O_2^+$ 398.2802; found 398.2804 $(M+H)^+$. IR (film, cm^{-1}) 3244, 3028, 2963, 2928, 2878, 2361, 1667, 1620, 1555.

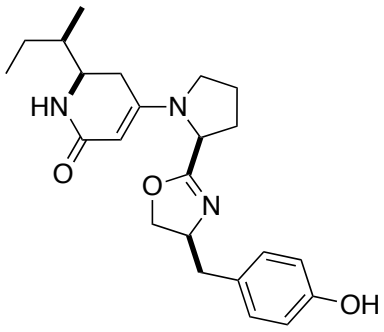
^{1}H NMR spectrum:



^{13}C NMR spectrum:

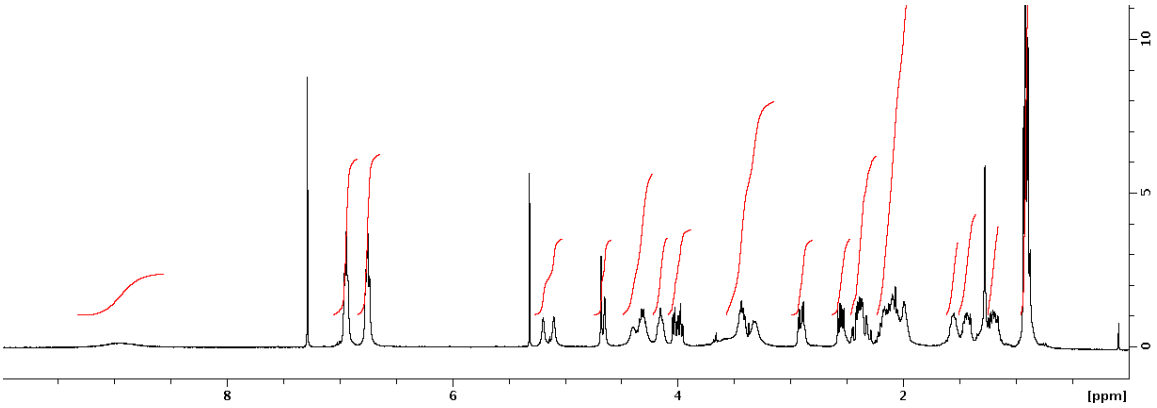


(*R*)-6-((*R*)-*sec*-butyl)-4-((*S*)-2-((*S*)-4-(4-hydroxybenzyl)-4,5-dihydrooxazol-2-yl)pyrrolidin-1-yl)-5,6-dihydropyridin-2(1*H*)-one (L,L,L-**30ipy**)

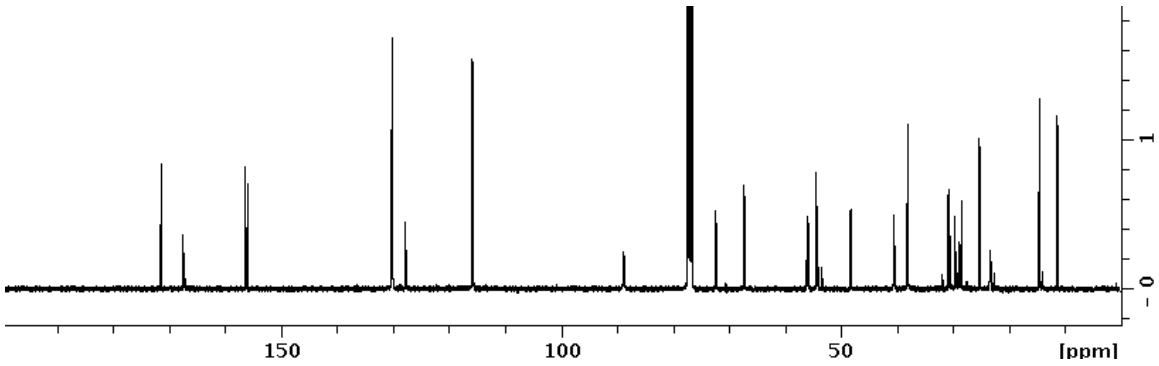


To a solution of compound L,L,L-**30ipy** in dichloromethane at 0 ºC, 5 eq of TMSOTf was added and stirred for 10 min. Solvent was reduced by rotavap and purified by flash column chromatography (3% MeOH/CH_2Cl_2) to obtain product as colorless oil. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.96 (br, 1H), 6.94 (t, *J* = 13.4 Hz, 2H), 6.75 (t, *J* = 13.4 Hz, 2H), 5.19-5.09 (br, 1H), 4.68-4.65 (br, 1H), 4.45-4.22 (m, 2H), 4.20-4.08 (m, 1H), 4.06-3.94 (m, 1H), 3.62-3.15 (m, 3H), 2.91 (dd, *J* = 16 Hz, 8 Hz, 1H), 2.55 (dd, *J* = 12 Hz, 8 Hz, 1H), 2.47-2.24 (m, 2H), 2.24-1.83 (m, 5H) 1.21 (m, 1H), 0.92 (m, 8H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 171.4, 167.4, 156.3, 156.0, 130.2, 130.0, 127.8, 115.8, 115.7, 88.8, 72.4, 67.3, 56.1, 54.4, 48.4, 40.6, 38.1, 30.8, 30.6, 29.7, 28.9, 28.5, 25.3, 23.4, 14.6, 11.4. HRMS (ESI) m/z calcd for $C_{23}H_{32}N_3O_2^+$ 398.2438; found 398.2429 $(M+H)^+$

^{1}H NMR spectrum:



^{13}C NMR spectrum:



pH Stability Assay

Procedure: Stock solutions of trimer DLD-**30vvv** and triphenylphosphine oxide (internal standard) in phosphate buffer (pH 6.4, 7.4 and 8.4) were prepared and stored in 25 ºC. Samples (25 µL) were injected into RP-HPLC system (C-18 column) at regular intervals.